Voice powered handsfree TV remote using Vesper's ZeroPower Listening™ Microphone

White Paper
Rev1.0

Author: Udaynag Pisipati

Vesper Technologies Inc.
77 Summer St, 8th floor
Boston, MA - 02110

WP3-Rev1.0
2/15/2019

Table of Contents

Introduction 4
The case of a TV remote 5
Are existing always listening solutions enough? 7
ZPL™ Microphones 9
How ZPL™ works 9
ZPL™ Design parameters 10
- Acoustic Threshold 10
- Hold Time 11

Advantages of ZPL™ in handsfree voice remote 11
Why Always-on, always listening handsfree remotes now? 12
Battery Life Analysis on a handsfree remote control: Analytical model 15
Design Changes compared to Push-to-talk remotes 17
- Microphone array 17
- Battery Size (mAh) 19
- Low-Power Wireless Technologies 20
- Other components 22

ZPL™ Design parameters for Handsfree TV remote: A Case Study 22
- Selecting the Acoustic Threshold 23
- Selecting the Hold time 25

Add-on benefits with Vesper's Piezo microphones 27
Conclusion 27
Citations 28

Vesper Technologies Inc.
77 Summer St, 8th floor
Boston, MA - 02110

Table of Figures

Figure 1: World's first wireless TV remote....................5
Figure 2: Flow chart for wake word detection using Wake on Sound mode....................8
Figure 3: Fixed adjusted WoS threshold, implemented with external resistor (R_g) between GA1 and GA2 pins....................10
Figure 4: Logged data from VM1010 (x-axis shows time in a 24-hr period)....................12
Figure 5: DSP Group DBMD6 Reference board with Vesper VM1010 Microphone....................13
Figure 6: Ambiq Apollo 2 Blue EVK (top) and Micro click board with Vesper microphones....................14
Figure 7: Energy depletion with Wake on Sound Vs alternate listening solutions....................16
Figure 8: Battery life savings with wake on sound Vs alternate listening solutions....................17
Figure 9: Battery Life savings (months) Vs Battery size (mAh) for a WoS solution....................20
Figure 10: Battery Life comparison with different communication protocols....................21
Figure 11: % Normal mode Vs. WoS Threshold....................24
Figure 12: Battery life (months) Vs. WoS Threshold....................25
Figure 13: % Normal mode Vs. Hold time (seconds)....................26
Figure 14: Far-field Voice Remote: Battery Life (months) vs. % Normal mode....................27

Vesper Technologies Inc.
77 Summer St, 8th floor
Boston, MA - 02110

Introduction

Always-on, always listening Voice control is becoming ubiquitous in modern day electronics. Voice interface does not require the push of a button or the user's full attention to activate the device. An on-board processor recognizes the wake word with a single microphone or a microphone array at the front end in order to stream the audio to the rest of the system. Recent advancements in hardware and software have led to the shipping of millions of smart speakers in the past couple of years. However, until now, the power-hungry nature of signal processing algorithms is limiting the devices to be wall powered as in the case of smart speakers or frequently plugged-in to the charger otherwise. In these devices, the microphone and speech processor are always alert, always waiting for the wake word - even when there is no sound to be heard. This leaves a lot of room for power optimization. How do we put them to sleep when there is nothing of interest to hear? How do we optimize the battery power consumption on these devices to turn every consumer's dream of using a true handsfree voice assistant into a reality – one that they could carry anywhere and everywhere while still obtaining a seamless user experience. Such a solution requires that every major component in the signal chain - from the microphone at the fore-front to the digital signal processor and the system on chip solution - to be ultra-low power and provide the best standby life possible on the device. For years, power optimization has been limited to the processor cores with the sensors at the forefront of the device still being a bottleneck for an ultra-low power design. Thanks to ZeroPower Listening™ (ZPL™) technology from Vesper Technologies [1], this dream of designing an end to end ultra-low power solution is finally becoming a reality, with energy harvesting starting from the microphone at the front end of the device. This white paper provides insight into the ZPL™ technology and emphasizes the importance of ultra-low power technologies to enable battery life savings in an always-listening system. An example of handsfree remote control will be considered as an analogy to demonstrate how ZPL™ enables new usecases for any other portable battery-operated devices including outdoor security cameras, smart hearables etc.

The case of a TV remote

Use of remote control has always been a challenging task since the invention of wireless remotes, perhaps attributed to the desire to keep the size of the remote minimal for the convenience of user. Look at the picture in Figure 1 below that shows the very first television wireless remote control, called Flash-Matic, built by Zenith technologies. The remote works by shining flash light on four photo cells on each corner of the television. Remote worked perfectly fine except on bright sunny days when the sunlight interfered with the photo cells causing interference in its operation. These interference and line of sight issues were later mitigated by Infrared and RF based remotes as the technology evolved.

Figure 1: World's first wireless TV remote

Now compare the lady operating the remote to control her TV in the picture to your personal experience in your own living room. On the bright side, you can easily visualize the evolution of the television industry both from a technological and programming perspective. However, you can notice that the user experience in operating the remote control to watch your favorite content has not changed much. Line of sight is not a major concern anymore, allowing the user to control the TV from anywhere in the room. However, the user still must hold the remote in his hand to control the television. Add to this the enormous content that is available for the user to search and consume on their television.

Push-to-talk (PTT) technology in the recent years has incorporated voice capabilities into remote control with the push of a button. Switching between various content on the TV is a very complex task to execute with only keypad control given the numerous viewing options available both on broadcast as well as IPTV platforms. Rapid industry adaption of PTT is itself an indicator of its search advantages compared to keyboard typing. But even a remote control with PTT still lacks the futuristic user experience offered by its counterpart – Smart TV. Moreover, the fundamental household challenge of finding the remote buried in the couch cushions and then discovering the specific program to view still prevails with a PTT design. Hands-free voice functionality, therefore, is an indispensable feature to add to TV remote control, whether you want to catch up on your favorite news channel while sipping a cup of coffee in the morning or watch a movie on a friday night lying on your couch with a bowl of popcorn.

A seamless user experience along with a battery life comparable to that provided by current push to talk solutions on a voice remote is the need of the hour for the TV industry. What if we can control and execute the commands on the TV using just our own voice to the remote. This next generation far-field voice remote presents a new level of user experience and ease of use. With a far-field remote, the user can say a wake word, such as Alexa, to wake up the TV remote located somewhere on a coffee table and issue a voice command immediately – no more searching for a lost remote and absolutely no hands involved. If you are closely following the voice interface euphoria in the press over the last two years, you might wonder why such a remote does not exist in the market already. Why is the design of such a power optimized handsfree remote so difficult for the TV industry to develop? What are the current technological constraints? Is there a solution or perhaps a perfect blend of solutions in the market that can help us achieve this design?

Answer to some of the above questions partially lies in the fact that consumers would not want to sacrifice the convenience of longer battery life with PTT remote for an enhanced user experience. Therefore, for a far-field voice remote to offer a valid user experience, it also must match the 6 months to 1-year battery life expectations of existing PTT solutions for normal everyday operation. This white paper aims to provide answers to the above questions in the context of Vesper's ZPL™ technology combined with other low power systems in the industry. The following section starts with existing always listening solutions and their associated challenges. Subsequent sections discuss ultra-low-power microphone solution from Vesper along with design changes that could help achieve the goal of building an ultra-low power handsfree voice remote.

Are existing always listening solutions enough?

Before we dive into ZPL™ technology, let's look at the challenges associated with building a hands-free voice remote. Imagine we want to replace the existing push to talk remote with an always listening solution that could use a capacitive MEMS microphone to build a far-field voice remote. In the traditional always listening solution, a voice activity detector will be on continuous lookout for voiced frames based on the speech activity level as shown in the large dashed box in Figure 2. Once a voiced frame is detected, it then triggers the rest of the system beginning with a wake word detection engine to identify if a wake word such as "Alexa" is spoken. If it discovers the voice frame to be a wake word, the system then transmits the command over the air through a low energy transport protocol such as Bluetooth Low Energy, Zigbee etc.

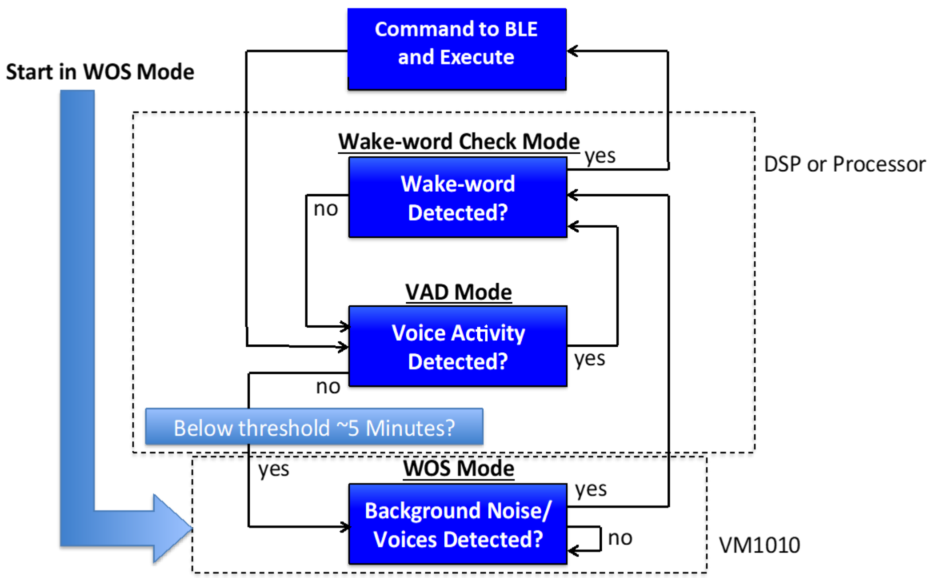

Figure 2: Flow chart for wake word detection using Wake on Sound mode

Microphone at the front-end interface of such a system will be operating in standby mode with a typical current rating of 200 μA on average for a capacitive MEMS product. This results in a total system standby power in the order of milliamps, from Voice Activity Detect (VAD) to wake word detection and command execution. Add to this the need for a microphone array to achieve higher far-field accuracy, standby power increases by atleast 2x compared to a single mic solution. The wake word detection system therefore becomes the only bottleneck for a system with low power transport such as Bluetooth Low Energy that only consumes 0.1μA on standby. A push to talk system on the other hand is only active when the user presses the microphone button, thereby consuming very little overall power for command processing. For a far-field remote to operate with a battery life comparable to push to talk solution, we need an always-on always listening system with an ultra-low power wake word detection at the front end followed by a system on chip solution that is on par with the front-end voice interface. Vesper's ZPL™ chimes in as a perfect solution for the wake word detection system with a simple, yet powerful wakeup system at the front end.

ZPL™ Microphones

ZPL™ is a new power optimized architecture for always-on listening systems with an ultra-low power sound detector. A ZPL™ microphone works by continuously monitoring the sound activity within the environment to match a specific threshold and frequency range within a certain sound level in the surrounding. While it's in this Wake on Sound (WoS) mode, the microphone consumes only 10μA of current and listens for sound activity in the range 65 – 89 dBSPL(A) within the voice band. Once it hears a sound that matches these characteristics, the microphone goes into a normal mode to stream audio to the system where it consumes 85 μA. Going back to Figure 2 above, the additional Wake on Sound mode in VM1010 acts as an acoustic watchdog to wake up the rest of the processor when there is a sound activity in the environment. A single VM1010 microphone can be used to trigger the microphone array or the DSP following the array to wake up from standby for command processing. In other words, ZPL™ integrated in VM1010 offers an ultra-low power mode even before the lowest power voice activity detect mode in the existing always listening solutions, thereby offering considerable power savings.

How ZPL™ works

Any piezoelectric MEMS microphone from Vesper works by converting mechanical energy into electrical voltage required to drive the circuitry from within. When a soundwave hits the piezoelectric MEMS element in the microphone and makes it move, the motion creates a voltage that is monitored by a very low power current comparator circuit. This comparator in turn sends a wake signal to the processor when the sound level exceeds a certain threshold level. This threshold in the range 65-89 dBSPL can be selected by a proper choice of the resistor to optimize the performance of VM1010 based on the surrounding environments. In addition, the microphone uses a bandpass frequency response from 250Hz to 6kHz to reduce false positives by responding to human voice, while rejecting wind noise, HVAC sounds and other environmental sounds. The default WoS threshold is 65dBSPL and can adjusted by connecting a resistor between pins GA1 and GA2 as shown in Figure 3. The smaller the resistor between GA1 and GA2, the higher the gain of the instrumentation amplifier. These pins provide access to the feedback network of an instrumentation amplifier in the WoS signal path.

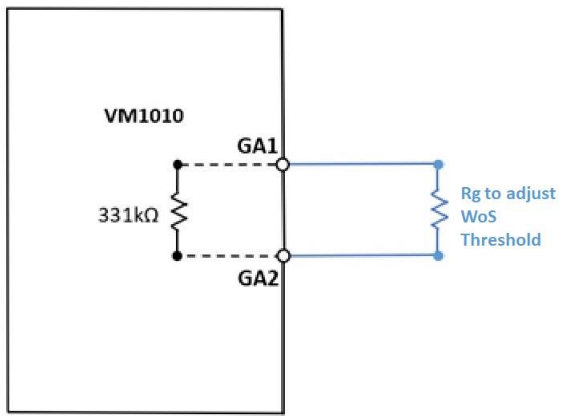

Figure 3: Fixed adjusted WoS threshold, implemented with external resistor (R_g) between GA1 and GA2 pins

ZPL™ Design parameters

Acoustic Threshold

ZPL™ uses a peak detect mechanism to trigger the microphone to ambient sound levels in the range 65 – 89 dBSPL. This wide range of acoustic threshold provides developers with the flexibility to fine tune the listening level of the microphone against the background environment. ZPL™ enables usecases for wake word activation as well as event detection and therefore the choice of acoustic threshold depends on the actual usecase. For a device with wake word activation, choice of threshold depends on the typical speech level of the user relative to the background noise level and the proximity of the user to the device. For devices without wake word activation such as a security camera sensing for activity in the outdoor environment, a higher acoustic threshold would be appropriate to optimize the battery life on the device, given the typical high background noise levels in the outdoor. Once a WoS threshold is chosen, the metric can be configured on the device using an external resistor.

Hold Time

Hold time is the time in which ZPL™ microphone switches back to listening mode from a normal mode after waking up to a sound activity. This metric needs to be programmed at the system level to optimize the battery life on the device. A system with short hold time improves the battery life by putting the microphone in sleep mode more often. On the other hand, a system with long hold time means that the microphone would already be out of the wake on sound mode when the wake word is spoken.

Advantages of ZPL™ in handsfree voice remote

How does ZPL™ enable battery savings compared to VAD only based always listening solution. A VAD only system continuously monitors for voice activity 24 hours in a day, out of which a vast amount of time includes no activity, whatsoever, in a living room environment – for example 6-8 hours of time when everyone is sleeping. Consider Figure 4 which shows the activity in a living room in a 24-hr period collected with a Wake on Sound microphone set to an acoustic threshold of 78 dBSPL. The x-axis shows the time in a day starting with mid night. The 1's on the graph indicate mode switch from Wake on Sound to Normal mode. The graph only shows activity in a single household in a single day, but it is obvious that the switching between wake on sound and normal mode happens only during the most active periods of the day and the microphone is in deep sleep majority of the day. For a TV remote, this means that the rest of the processing units on the remote including A/D converter, voice processor and the main processor can be in standby mode for most of the time in a day. This selective triggering with WoS mode therefore saves standby power compared to a VAD only solution.

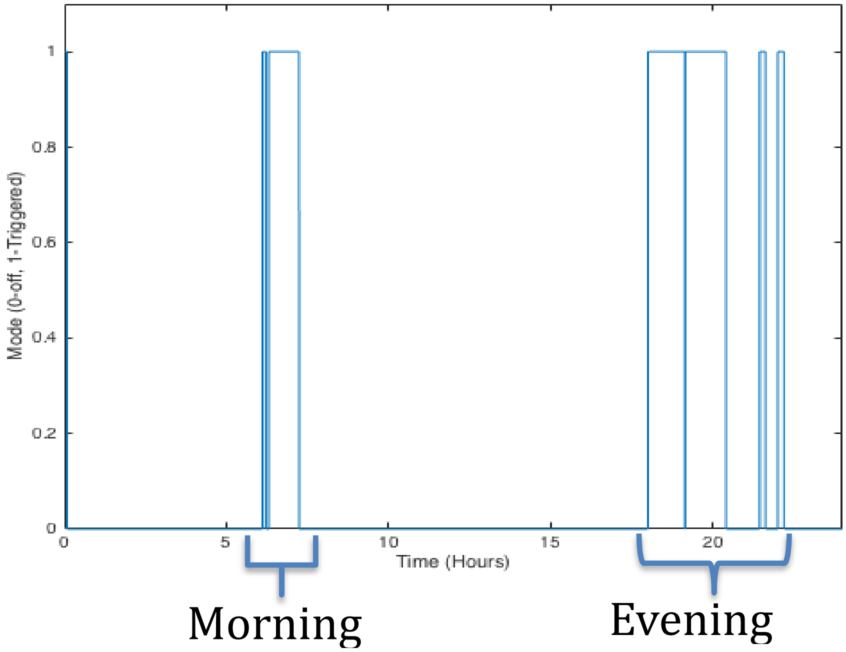

Figure 4: Logged data from VM1010 (x-axis shows time in a 24-hr period)

Why Always-on, always listening handsfree remotes now?

While ZPL™ provides a novel microphone architecture suitable for portable devices, power optimization can only be achieved with an end to end system that can operate in ultra-low power mode. Therefore, it's worth considering what changed in the industry in the last few years that can lead us in the path to experience this frictionless user interface for voice remotes. A handsfree voice remote will have to cancel the background noise in the environment for an accurate wake word detection and then transmit the processed user command to the TV using an ultra-low power transport protocol. The power constraints must be minimized at several layers of the far-field remote design to achieve an overall battery life savings on the remote. Starting from the microphone to voice processor to the system on chip used for transport, all the underlying components must operate on ultra-low power to achieve extended battery life. For a voice processor, there are several ultra-low power alternatives emerging in the market both for analog and digital interfaces. DSP Group DBMD4P Audio/Voice SoC mentioned in above sections, utilizes its proprietary HDClear™ voice technology to enable background noise suppression and Acoustic echo cancellation for improved far-field accuracy [2]. DBMD4P, as shown in Figure 5, offers analog and digital interfaces with a low power mode which consumes ~40µW in standby and only ~3mW when in full processing mode

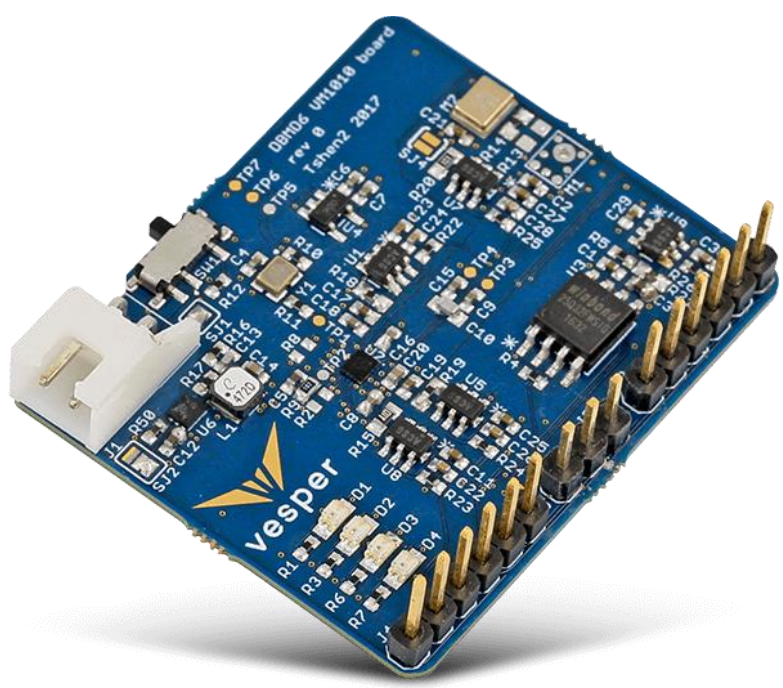

Figure 5: DSP Group DBMD6 Reference board with Vesper VM1010 Microphone

Ambiq Micro's Apollo line of products [3] is another example of processors designed to handle highly computational audio signal processing tasks in portable devices. The ARM Cortex-M4F based Apollo series offers multi-mic array signal processing with the highest energy efficiency based on their SPOT™ (Subthreshold Power Optimized Technology) platform.

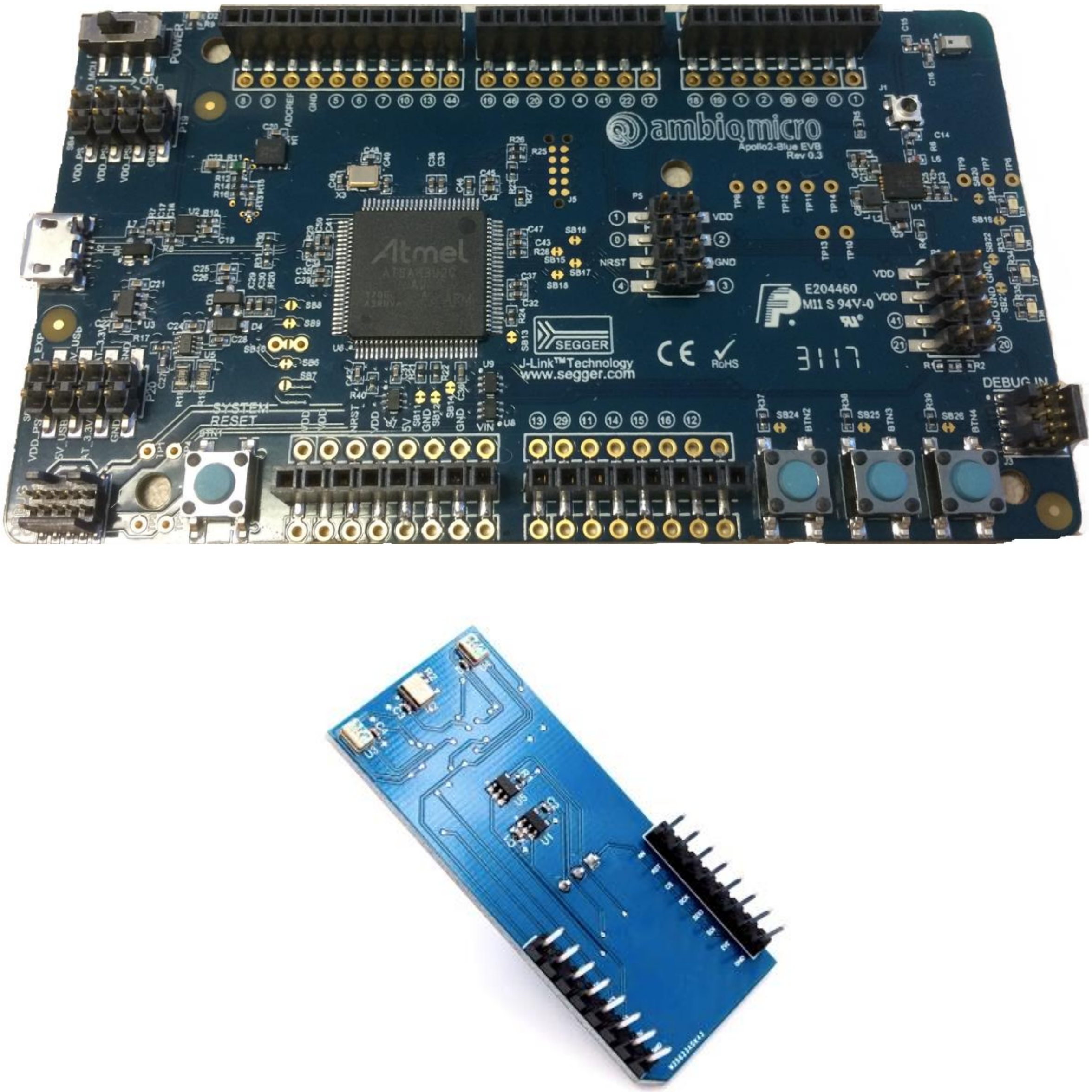

Figure 6: Ambiq Apollo 2 Blue EVK (top) and Micro click board with Vesper microphones

The Apollo Blue series further integrates a Bluetooth Low Energy radio providing power savings that are not possible with alternate MCU solutions thereby making it appropriate for handsfree voice remotes. For battery-powered operation, the Apollo2 Blue and Apollo3 Blue achieve ultra-low power processing with just 10 μA/MHz and 6 μA/MHz respectively in full power mode [2]. Vesper's ZPL™ technology, together with these SPOT-based ultra-low power chips mark the beginning of a new era of portable electronic devices. Apart from handsfree voice remotes, these technologies also enable multiple use cases for battery operated products including security cameras, hearables and wearable devices. A reference design with Ambiq Apollo 2 Blue and Vesper microphones is shown in Figure 6.

Battery Life Analysis on a handsfree remote control: Analytical model

At Vesper, we developed an analytical model to analyze the battery life savings with a Wake on Sound microphone compared to alternate listening solutions. This section provides the background of the analytical model while the following sections describe the impact of ZPL™ metrics such as WoS Threshold and hold time on the battery life savings. The analytical model takes as input the duration of the events, processing times and power consumptions for the active components in the system for the specific usecase and provides a comparison of battery life with and without Wake on Sound technology. It is assumed that the remote control is powered by 2 AA batteries at 1.5v and 2400 mAh capacity each. A VM1010 microphone combined with 2 additional digital microphones are considered in the system to enable far-field processing on the DSP. The model uses the current consumption measured on VM1010 microphone and DSP Group DBDM6 platform with Sensory wake word detection engine to calculate the total power consumption for the various modes - Wake on Sound mode, voice activity detect mode, wake word processing and full power mode - in an always-listening system shown in Figure 2. To provide a baseline, the calculations in this section are performed under the assumption that the VM1010 is set to 65 dB Threshold, which is the lowest acoustic threshold setting on the microphone. Following sections describe the impact of using higher thresholds on the battery life savings.

Figure 7 shows the charge depleted by a TV remote using ZPL^{TM} technology compared to alternate listening solutions.

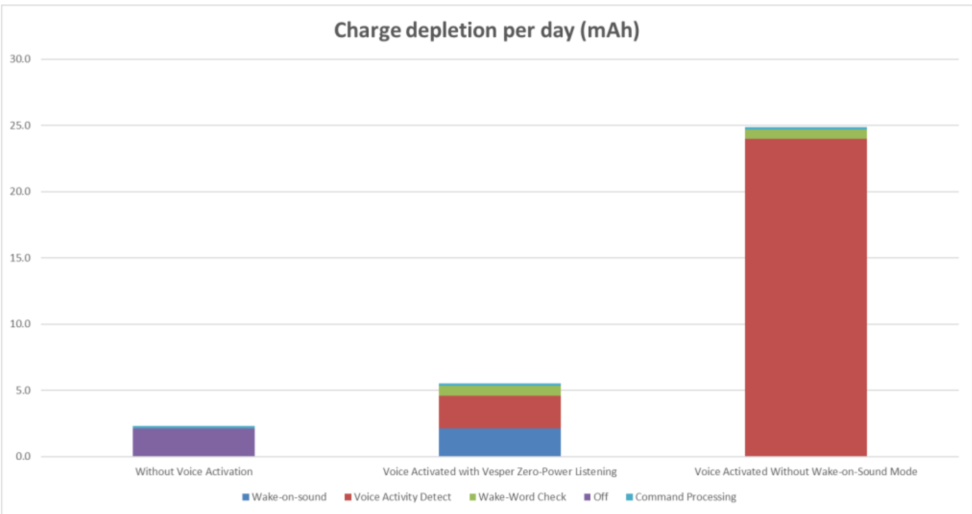

Figure 7: Energy depletion with Wake on Sound Vs alternate listening solutions

It can be observed that voice activity detection mode in a solution without Wake on Sound microphone consumes more than 80% of the battery capacity at any given moment. By adding a Wake on Sound mode before the voice activity detection, VM1010 microphone optimizes the time that the system goes into voice activity detect mode. These savings in energy directly translate into battery life savings on the voice powered remote as shown in Figure 8. WoS microphone increases the standby life of remotes by 10x and provides an overall battery life savings 4x times that of alternate listening solutions with a typical daily use. On the other hand, battery life savings from a WoS solution are also comparable to push to talk solution given the enhanced user experience with a handsfree remote. These savings can further be improved by optimizing the Wake on Sound threshold and power consumption on the DSP board to meet the needs of remote control usecase.

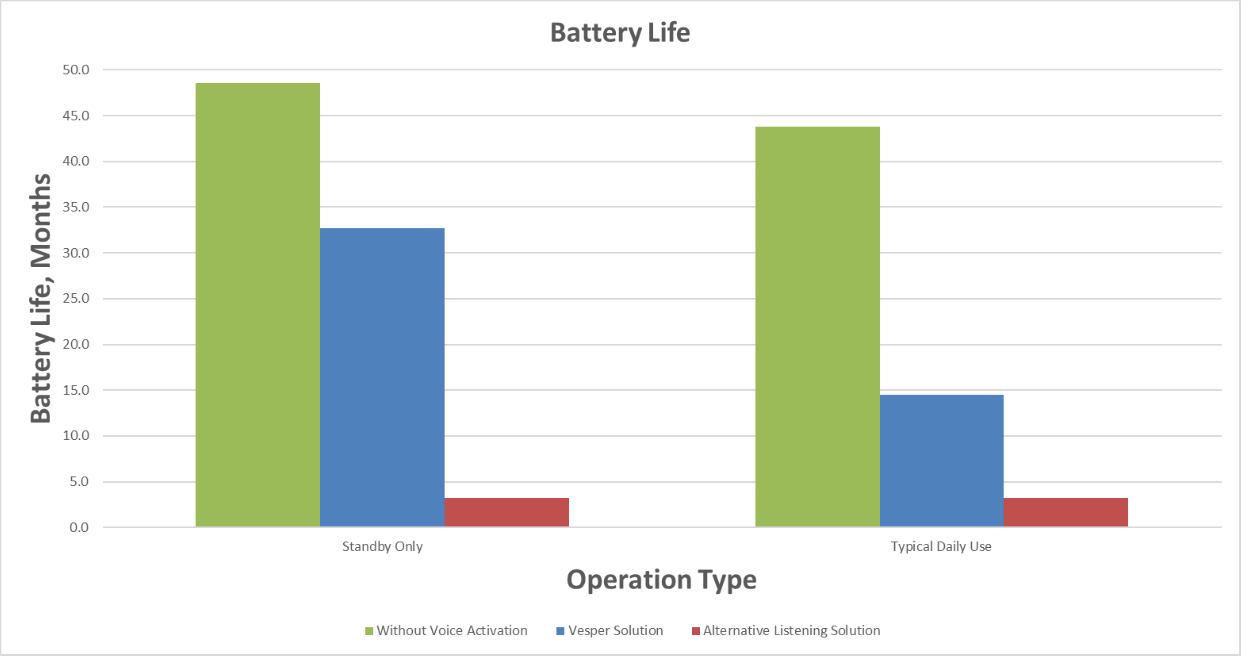

Figure 8: Battery life savings with wake on sound Vs alternate listening solutions

The following sections utilize the above model to show the impact of various design changes required for a hands-free TV remote on the battery life of the device.

Design Changes compared to Push-to-talk remotes

Microphone array

Push-to-talk remotes use a single microphone since the user is holding the microphone close to their mouth with a button press to talk to the television. In a handsfree scenario, the remote is usually sitting on a coffee table in a living room and the user would be talking to the remote in any direction and at any angle. Voice wakeup also requires the use of a wake word to wake up the system first and then utter the voice command. Use of a multiple microphone array enables higher signal to noise ratio on the system and improves the wake word detection and response accuracy of any far-field device such as a handsfree remote control. In a system with ZPL™, a VM1010 microphone can be coupled with one or more microphones that would be primarily used for noise suppression. VM1010 would then work as an acoustic watchdog to wake up the system if the activity level in the living room exceeds the set threshold. Audio stream is then passed to the signal processor for voice activity and wakeword detection. Our case study above assumes a VM1010 microphone coupled with 2 digital microphones. The additional microphones would be required if the system uses beamforming algorithms for noise suppression. For a system without beamforming and operating on a single mic noise reduction, one additional microphone would be enough.

Contrary to expectation, the integration of VM1010 microphone into an existing remote-control design does not need a huge design change on the voice processor. Compared to capacitive MEMS microphones, use of VM1010 only requires two additional GPIO control pins on the processor and an external resistor on the PCB to control the threshold setting. The two GPIO pins are connected to the Mode and Dout on the VM1010 to switch the microphone between WoS and Normal modes. While in WoS mode, the microphone will output a digital high through the Dout pin when a sound is detected. The processor then sets the Mode pin to digital low to switch the microphone to Normal mode. When the user event is completed, processor can then set the Mode pin to digital high to switch back to WoS mode. Selection of design parameters for ZPL™ are discussed in the following sections.

Battery Size (mAh)

Standard TV remotes are generally powered by either 2 AA or 2 AAA alkaline batteries at 1.5V each. The expected battery life of the remote depends on the battery size defined in milli-amp hours. While AA batteries are a traditional industry standard, for example, in an xFinity TV remote, some designs use AAA batteries to achieve smaller form factor. While battery run time on AA batteries is twice that of AAA batteries with half the capacity, a smaller size battery achieves a smaller form factor design. A few other gadgets such as Apple TV remote or Logitech Harmony Elite Remote come with a rechargeable Lithium ion battery such as CR2032 operating at 1050 mAh at 3.7V. Even these devices with rechargeable batteries need to last atleast a couple of weeks on a single charge as consumers often leave them on the couch instead of returning them to the charging base daily. In all these cases, Wake on Sound mode reduces the need to replace batteries more often making them last for months and even upto an year. Alternately, VM1010 microphone when combined with ultra-low power digital SOCs mentioned in the earlier sections enable the use of a coin cell battery and therefore ultra-compact form factor for handsfree remotes while still providing optimum battery life. Figure 9 shows the impact of Battery size on the battery life provided by Wake on Sound mode. For any given battery size, WoS mode offers better battery life performance compared to alternate listening solutions even when there is 5 hours TV playback in the background. Impact of ZPL™ microphone is only 2% of battery life from a typical use but adds always listening capabilities to the TV remote.

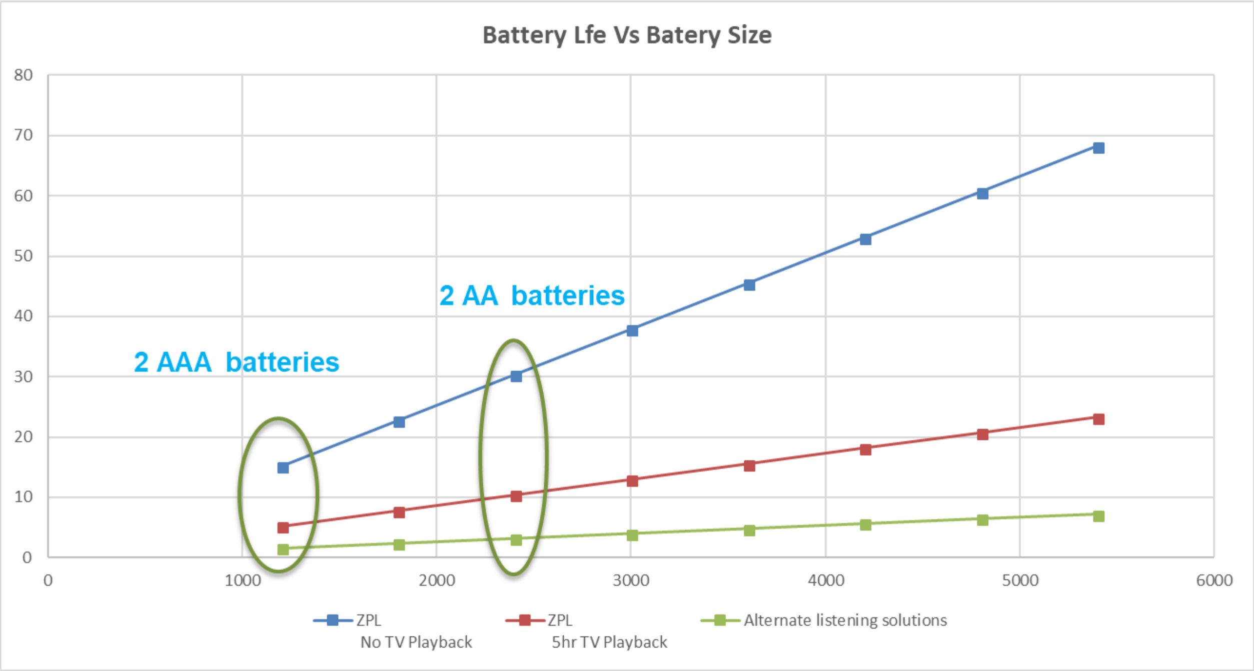

Figure 9: Battery Life savings (months) Vs Battery size (mAh) for a WoS solution

Low-Power Wireless Technologies

Ultra-low power communication technologies such as Wi-Fi, IrDA, Zigbee RF4CE and Bluetooth Low Energy (BLE) offer energy harvesting with very low duty cycles such that they can stay in sleep mode for extended periods of time. Power savings from Wake on Sound mode are agnostic to the low power transmission technology used in the voice remotes but choosing a transport protocol with ultra-low power standby current and minimal latency is the key to optimize battery life on a TV remote while providing a seamless user experience. Voice remote will be in full power mode only during the command processing time, which is only a fraction of TV viewing time in a typical household. Therefore, to enable seamless response from the handsfree system, the wireless technology must also be chosen to provide minimum latency for voice processing along with optimum battery life. Table 1 below shows a comparison of different technologies [4], their power consumption and latency. A comparison of battery life for each of these technologies is also shown in Figure 10. Zigbee RF4CE (Radio frequency for Consumer Electronics) replaced the Infrared technology over the years because of its primary advantage that the remote does not have to be in the line of sight to control the television.

Parameter / Protocol	BLE	RF4CE	IrDA	Wi-Fi
Standby current (mA)	0.0001	0.0018	N/A	10
Tx Current (mA)	8	30	1.95	100
Datarate (bps)	960	250k	121	40M
latency (msec)	2.5	20	25	1.5

Table 1: Comparison of different Low Energy communication protocols

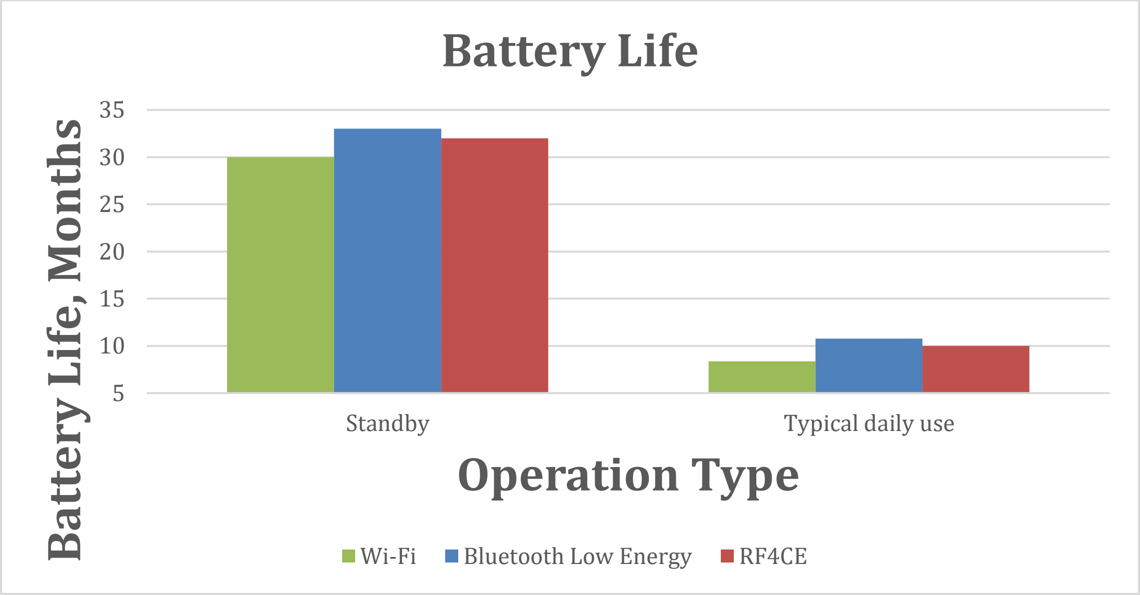

Figure 10: Battery Life comparison with different communication protocols

Bluetooth Low Energy provides the optimum current consumptions out of the four different technologies and offers minimum latency of 2.5 milliseconds which is crucial for a frictionless experience with far-field remote control. Ultra-fast wakeup time of 200 µSec on VM1010 combined with low latency SoC processors and BLE technology enable a seamless voice experience on a handsfree TV remote. ZPL™ enables this capability at lowest power levels when used in conjunction with ultra-low power wireless technologies.

Other components

Some PTT remotes also use an Accelerometer to wake up the DSP when the user lifts the handset to press the microphone button. A handsfree remote with VM1010 could potentially eliminate the need to have an accelerometer used for lift-up. Accelerometers usually have an ultra-low current consumption in the order of µA, however removing the component could result in BOM savings in the system design and also cost savings for remote manufacturers.

ZPL™ Design parameters for Handsfree TV remote: A Case Study

In order to study the frequency of wake on sound triggers in a living room environment in normal households, Vesper designed data loggers with VM1010 microphone at different threshold levels – 65 dB, 71 dB and 77 dB. These data loggers were then distributed to the employees to record data in their living rooms during a 24 hours period. In a typical living room environment where handsfree voice remote will be used, the user is located at 1-2 meters from the remote. The voice remote in turn lies on a coffee table another 2 meters from the TV. Therefore, the dataloggers were located on a coffee table within the living room. Considering that the activity in a living room is different from a weekday vs a weekend, the experiment is repeated for both scenarios. The WoS triggers from the data logger is then analyzed with hold times varying form 7 seconds upto 5 minutes. This section analyzes the impact of two design metrics - Acoustic Threshold and the hold time - on the battery life performance based on 700 hours of data recorded from 24 different households. These households ranged from 1-5 people in the household and included pets, kids, birds etc.

Selecting the Acoustic Threshold

An average US household watches 4 hours of television in a typical day. Given this information, it might be tempting to think that the Acoustic threshold on the remote need to be set to a higher SPL value to reduce false wakeup. While it is true that the VM1010 will be triggered a few times while the TV is on for those 5 hours. But the peak detect mechanism on VM1010 minimizes the false wakeups from the TV programming. The typical average volume settings used on the television and the distance of the voice remote from the TV would automatically compensate for the peak sound pressure level at the WoS microphone of the remote. Our household data which includes false wakeups from the active Television time as well as any wakeups from other background activity in the living room shown in Figure 11 indicates that for any given acoustic threshold, the normal mode time falls only between 15-25% of the time in a day. On the other hand, in a quiet living room without the television powered on, the background noise in the living room falls below the 65-dB minimum threshold setting available on VM1010. Therefore, VM1010 would be in WoS mode and listening to the sound activity for most of the time in a 24-hour period. On the other hand, the typical user speech levels at the ZPL™ microphone on the coffee table would be in the low 70 dBSPL. Setting the acoustic threshold at higher levels might require the user to shout at the remote thereby compromising on the user experience. In addition, our data shows that the savings in normal mode time and hence the battery life as shown in Figure 12 are only a few percent at higher threshold such as 71 dB compared to the 65 dB minimum acoustic threshold setting. Therefore, for a handsfree TV remote, 65 dB acoustic threshold is recommended setting that provides the tradeoff between false wakeups and optimal battery life.

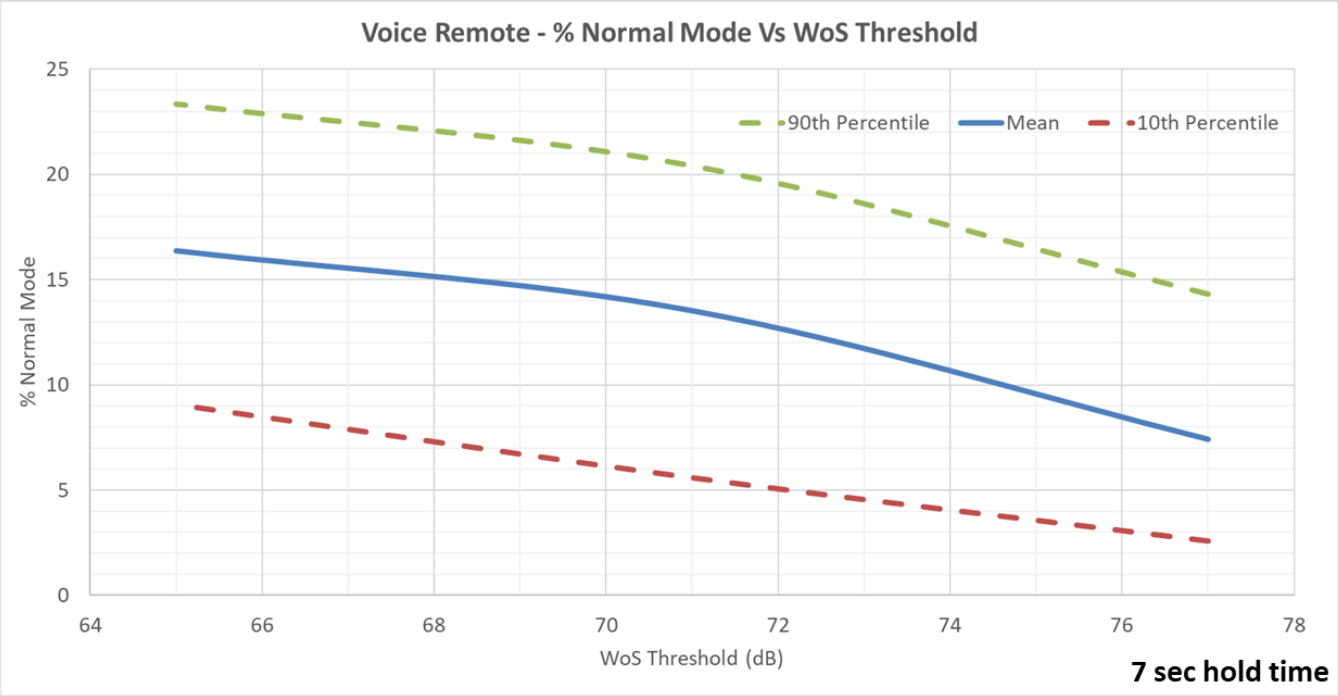

Figure 11: % Normal mode Vs. WoS Threshold

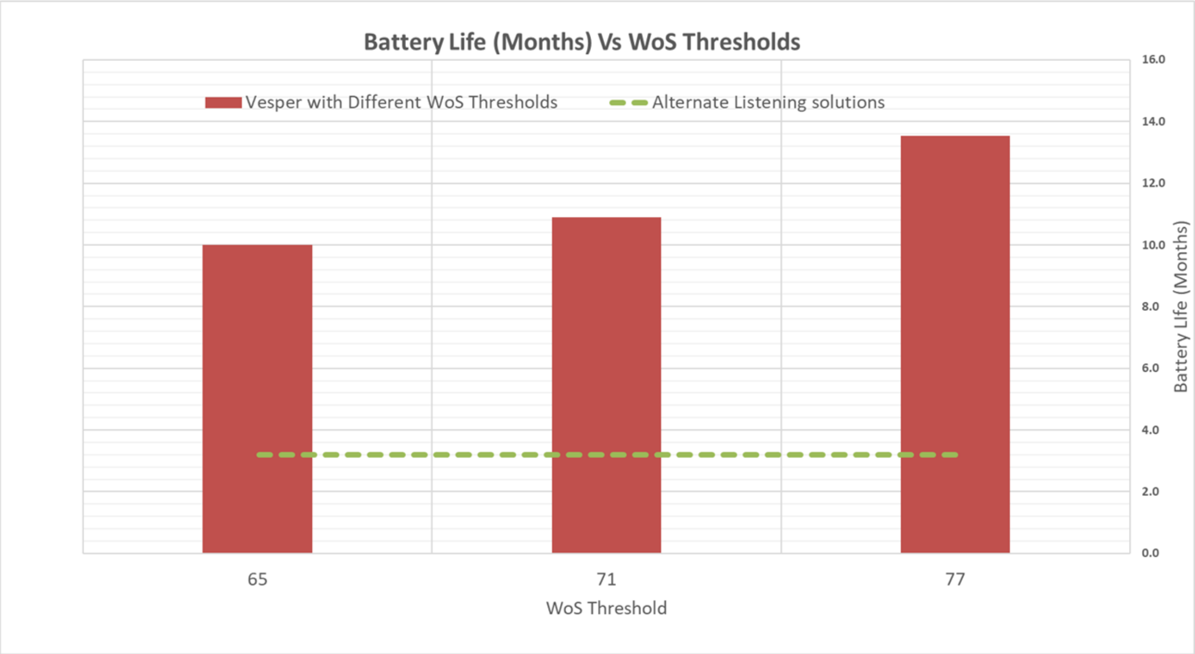

Figure 12: Battery life (months) Vs. WoS Threshold

Selecting the Hold time

Hold time is the time in which the VM1010 microphone goes back to WoS mode after a normal mode activity. This metric must be programmed at the system level, so the DSP can monitor the activity in the background and put the device back to sleep if needed. Figure 13 shows the amount of time VM1010 would be in normal mode at hold times from 7 seconds to 5 minutes for a given 65 dB acoustic threshold setting. 7 second hold time reduces the normal mode activity by half compared to 5-minute hold time. Programming the DSP for a shorter hold time improves the battery life on the device since VM1010 would go back to WoS mode more often. Selecting a shorter hold time is also beneficial for living room environments where there is background speech activity that could result in false wakeup. In this case, choosing a longer hold time could penalize the system heavily for a false wakeup. On the other hand, care should be taken to have the microphone in normal mode for atleast the amount of time a required voice command would be uttered. In a handsfree voice remote where the actual voice command such as “Play”, “Pause”, “Open House of Cards” combined with the wakeword is in the order of a few seconds and a hold time in the range 7 sec to 30 seconds is recommended.

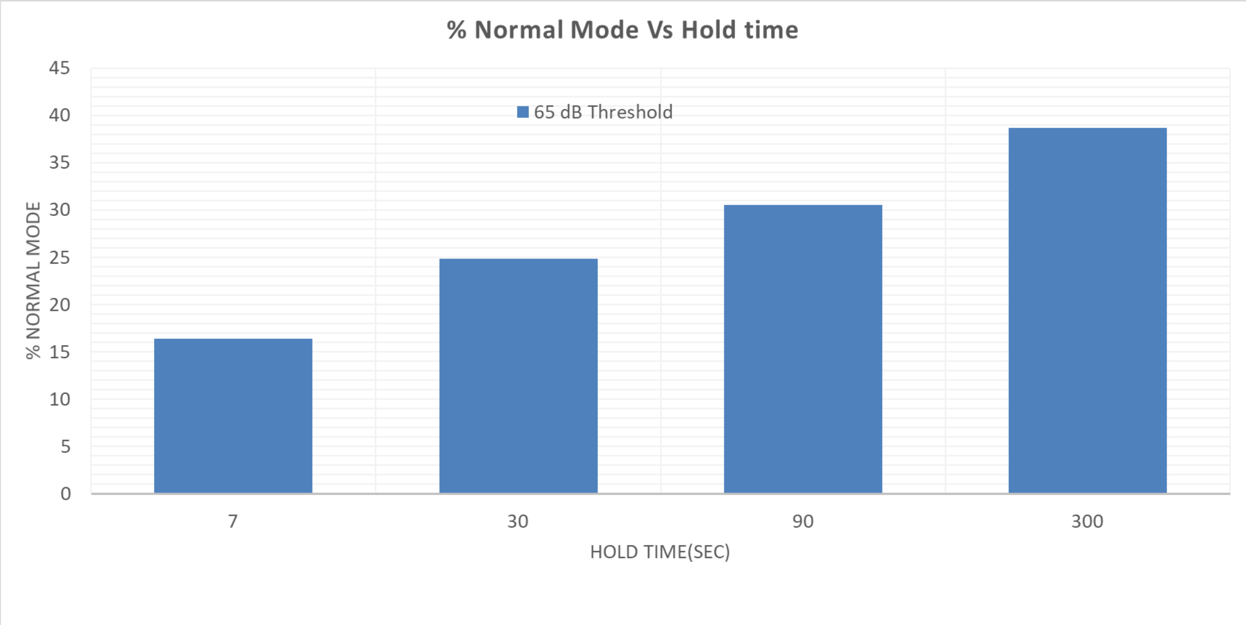

Figure 13: % Normal mode Vs. Hold time (seconds)

Graph in the Figure 14 summarizes the battery life savings from WoS microphone for a handsfree TV remote for the recommended settings on VM1010. ZPL™ provides 10x increase in standby battery life. For typical battery life time, setting the VM1010 for the minimum recommended settings of 65 dB for acoustic threshold and 7 second hold time provides a 3x battery life advantage compared to alternate listening solution, However, battery life upto 7x can be achieved using a higher acoustic threshold setting.

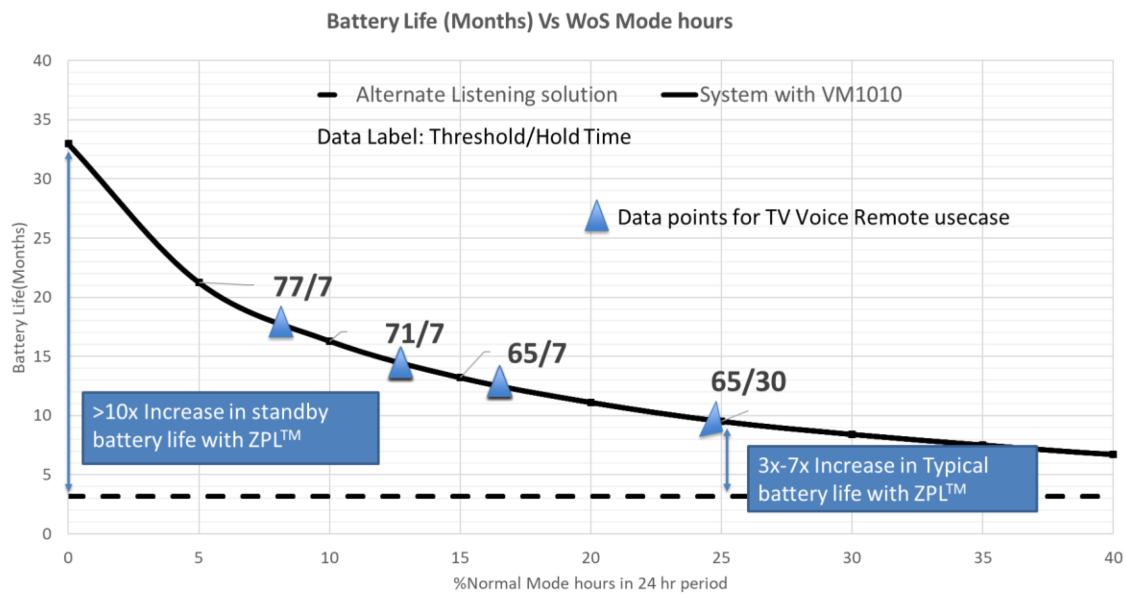

Figure 14: Far-field Voice Remote: Battery Life (months) vs. % Normal mode

Add-on benefits with Vesper's Piezo microphones

Vesper's piezoelectric microphones also offer additional advantages for the voice remote usecase. Piezoelectric MEMS microphones have a quick startup time of 200 µsec which is 1000x less than a capacitive MEMS microphones enabling higher keyword detection accuracy. Piezoelectric material is inherently robust to environmental contaminants such as water, dust and even kitchen oil or popcorn butter, thereby offering robust performance for the long term.

Conclusion

In a future that is powered by voice, Vesper's wake on sound technology helps to provide an optimal battery life and a friction less interaction for consumers to interact with their smart TVs from across the room. ZPL™ combined with ultra-low power audio SoCs enable an ultra-low form factor far-field voice remote that could be operated with coin cell batteries while still offering the convenience of a Push-to-talk system with extended battery life. All these benefits come with the minimal design change required to implement ZPL™ into existing remote designs. With the barriers of power consumption eliminated from the remote-control design, one can only imagine more and more Smart TV manufacturers adapting voice remotes with Vesper microphones.

For additional details on Voice remote case study and Vesper's product portfolio, please reach out to info@vespermems.com

Citations

[1] "Vesper Technologies," [Online]. Available: https://www.vespermems.com.

[2] "DSP Group," [Online]. Available: https://www.dspg.com/products/ .

[3] "www.ambiqmicro.com," [Online]. Available: https://www.ambiqmicro.com.

[4] "Digikey," [Online]. Available: https://www.digikey.com/en/articles/techzone/2017/oct/comparing-low-power-wireless-technologies.