2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
October 11-14, 2020. Toronto, Canada

HDR (High Dynamic Range) Audio Wearable and Its Performance Visualization

Danson Evan Garcia, Alexander Mertens, Jesse Hernandez, Mei Li, Steve Mann
MannLab Canada, 330 Dundas Street West, Toronto, Ontario, M5T 1G5

***Abstract*—We introduce a novel use of wearable technology to reconstruct and output high dynamic range (HDR) audio signals. The purpose of designing an end-to-end HDR audio wearable device is to create an audio device capable of replicating and even extending the dynamic range of human hearing. We also introduce a different method of visualizing audio by way of a rotary Sequential Wave Imprinting Machine (SWIM). The results show that real-time HDR audio processing and visualization are possible at the standard audio sampling rate of 44.1 kHz. Additionally, the quality of the recorded audio improved from the originally captured data after HDR processing. We further analyze the dependence of different audio gain combinations as well as the influence of the number of audio input channels on the quality of the HDR signal reconstruction in various acoustic environments.**

***Index Terms*—Signal visualization, signal processing, audio signal, signal reconstruction, high dynamic range (HDR).**

Fig. 1: The wearable HDR audio device worn by a user.

I. Background and Introduction

Physically measured signals that are converted into a digital form may experience saturation due to the inherent limitations of the receiver's dynamic range. In audio, microphones have fixed capabilities for detecting audible signals. Signal capture is performed by an analog-to-digital converter (ADC), which digitizes the incoming analog audio signal into discrete values. At each sampled time, n, the ideal relationship between the analog input, $x[n]$, and digitized output, $y[n]$ is easily modeled using the piece-wise transformation in Eqn. 1. Q is the number of ADC bits while V_{min} and V_{max} represent the voltage range of the conversion. $\lfloor\ \rfloor$ denotes a floor function. $V_{\text{p-p}}$ is the peak-to-peak voltage range of the ADC. Prior to digitization, the analog input may be amplified using a gain stage circuit.

$$y[n] = \begin{cases} \left\lfloor \frac{x[n]}{V_{\text{p-p}}} \times (2^Q - 1) \right\rfloor & \text{if } V_{\text{min}} \leq x[n] \leq V_{\text{max}} \\ 2^Q - 1 & \text{if } x[n] > V_{\text{max}} \\ 0 & \text{if } x[n] < V_{\text{min}} \end{cases} \quad (1)$$
$$V_{\text{p-p}} = V_{\text{max}} - V_{\text{min}}$$

Issues arise due to the information loss in the digitization process. If a change in the analog signal is smaller than the quantization step, then the change will not be recorded by the ADC. Also, if the amplitude of an incoming signal is too high, then the conversion will reach a saturation point. In other words, signals that are either too small or too loud are not fully captured by an ADC.

Methods using signal estimation in time and frequency domains on a single channel output signal have been proposed to compensate for the limitations of the receiver circuit used in the system [1]–[5]. However, real-time reconstruction relies on access to fast processing units, making them difficult to use on a wearable device. Faster methods exist where some knowledge of the input signal is used to reconstruct the signal efficiently [6]–[9]. These methods utilize one channel at a lower gain as a reference and another channel that maximizes the ADC dynamic range.

A. High Dynamic Range (HDR) Technique

Another method which utilizes multiple channel compositing is called the high dynamic range (HDR) technique, more widely and commercially used in imaging [10]–[14] and video [15]–[19].

The idea of HDR imaging is to capture images at multiple levels of exposure, and combining the obtained information to create a composite that mimics how our human eye perceives the real world. In effect, this collated image carries a greater dynamic range compared to a single image taken at a particular exposure level. In measured signal applications, HDR exploits the gain interleaving [8] and dynamic range extension [9] methods by using a multi-channel signal measurement system at different gain settings while utilizing all the signal information in the collation process. The individual signals are usually combined via the use of a certainty function which represents the certainty of a measurement. A weighted average utilizing the certainty values form the HDR composite [20], [21].

B. Sequential Wave Imprinting Machine (SWIM)

Sequential Wave Imprinting Machines were invented to visualize phenomena that exist in the surrounding environment that was previously invisible [22], [23]. The machine can be used

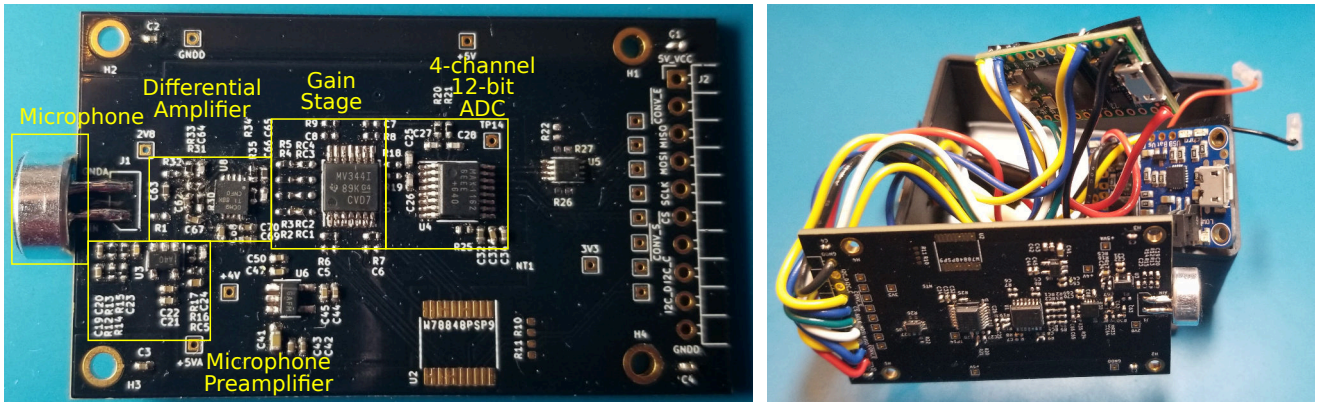


Fig. 2: Left: Front side view of the custom-built audio conditioning circuit. Right: The wearable HDR audio device with its internal components shown.

as an educational tool in classrooms to better understand the topic of waves [24], [25]. The visualization machines are able to make real otherwise abstract concepts, creating an augmented reality or extended reality [26], [27].

The device and visual overlay system work in a rotary setting [28], [29] which can be extended to audio signals to show volume amplitudes and saturation points of specific microphones. In the context of audio HDR, we have used this to visualize the audio signals and their ranges for saturation for a specific microphone. SWIM has made the final composited signal and the process of how it is created easier to visualize and understand.

C. Motivation

In audio, affordable recording devices are unable to capture the signal with a high dynamic range. Commonly used audio devices are able to reactively adjust the input gain in order to accommodate changing volumes in the soundscape. These methods are less able to respond to quickly changing volumes in a soundscape [8]. We propose a proactive method to achieve this that artificially increases the dynamic range of any ADC. In addition, signal clipping introduces additional high-frequency noise which leads to static and screeching sounds. The high-frequency noise may also cause the output system to be damaged via overheating. Clipping also has negative effects on automatic speech recognition performance [5].

This paper puts the concept of HDR audio into the design of an affordable wearable audio recording device that is able to capture an adjustable dynamic range of audio. As proof, this paper aims to show the increase in signal quality of an HDR composite over using a single channel for an arbitrary audio signal. The end goal of which is to create an audio device that can mimic and potentially extend the dynamic range of human hearing.

II. Wearable HDR Audio Design

To prove the concept of HDR compositing beyond images and videos, we have designed a working wearable HDR audio acquisition system that will serve as the test bench for our experiments. The overall system block diagram of the acquisition and streaming system is shown in Fig. 3.

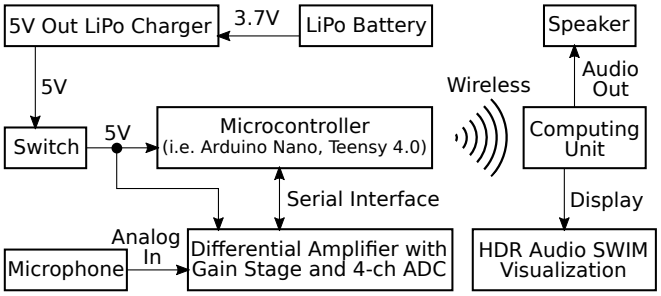


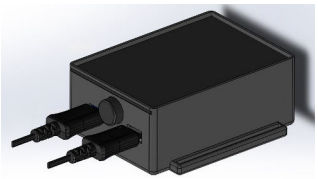
Fig. 3: System level diagram of the wearable HDR audio design.

A $3.7\,\mathrm{V}$ $1800\,\mathrm{mA\,h}$ lithium-ion polymer (LiPo) battery is used and a LiPo charger with a boost converter converts the battery's voltage into a constant $5\,\mathrm{V}$ supply. A rocker switch controls when the devices receive power. The audio analog input of the system is an electret microphone. A commercially available wearable microcontroller development kit is used to interface with the computer and the custom-built printed circuit board (PCB). Aside from interfacing, the signal compositing is handled by the microcontroller to show that the processing is simple and effective. Once the digital signals have been sampled and processed, the microcontroller can send the data either wirelessly or using a serial interface to a wearable computing unit such as a smartphone or tablet. Further processing is then done on the more powerful computing unit to visualize the audio and the HDR compositing while potentially, listening to the processed acoustic signal through the speaker or headset.

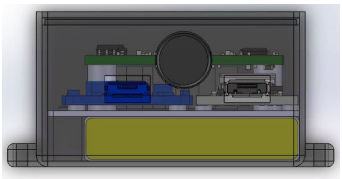
A. Mechanical Design

The box dimension of the design is $47\,\mathrm{mm} \times 43\,\mathrm{mm} \times 26.5\,\mathrm{mm}$. This is small enough to be worn on the wrist and can be seen in Fig. 1. The design prioritizes the device being worn on a daily basis while maintaining a compact and discrete form factor. The microphone is soldered on the circuit board to ensure the stability of the receiver while the device is in motion. The wearable device uses a stack design with three levels. As shown in Fig. 4, the battery forms the first level, the battery charger, microcontroller and the switch form the second, and the custom PCB with microphone forms the last. In Fig. 2, the final design with its internal components is shown. We use

standoffs to locate and securely mount the components and electronics on top of each other. The power switch and USB port holes for the battery charger and Arduino USB connectors are mounted to the wall of the case for easy access.



(a) Isometric view.



(b) Front view.

(c) Left side view.

Fig. 4: 3D model of the device. a) is the isometric view of the device with attached Universal Serial Bus (USB) cables. b) and c) are the front and left side view of the device with hidden details shown.

B. Hardware Design

The custom PCB consists of three amplifier stages and an ADC stage as shown in Fig. 5. As a point of reference these stages are highlighted on the actual PCB in Fig. 2. By having a custom PCB, we can easily modify key parameters such as the gain settings and filter transfer functions. This ensures that the signals we are gathering are phase coherent and have the proper gain settings for our experiments.

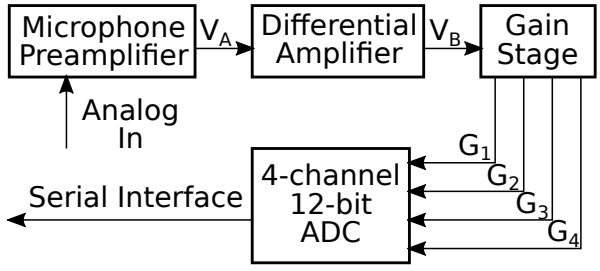


Fig. 5: A simplified block diagram of the custom-built differential amplifier with gain stage and ADC.

The microphone preamplifier stage is an input conditioning circuit to boost the weak analog signal coming from the microphone to the required signal line level. The circuit described in Fig. 2 of the MAX4465 device [30] datasheet is the circuit topology that we used. This circuit ensures that the analog input frequency range is within $20\,\mathrm{Hz}$ to $20\,\mathrm{kHz}$ while boosting its signal by $11\times$.

The differential amplifier stage takes the output voltage from the preamplifier stage, V_A, and subtracts it to a reference voltage. At this stage, we increase the bandpass filter roll-off by setting appropriate resistor and capacitor values to ensure most of the unwanted signals outside of the audio range is removed. The amplifiers are set at unity gain since the gain stages will handle the gain conditioning.

The gain stage consists of four opamps in inverting configuration. Each opamp is connected to a reference voltage and the signal input as the voltage output from the differential amplifier stage, V_B. The output of the opamps corresponds to the values $\{G_1...G_4\}$. The gains are easily set by changing the resistor values on the PCB. The default gains on the board are set at $1\times$, $4\times$, $16\times$, and $32\times$.

The ADC stage digitizes the 4 parallel gained voltages and sends these values to the microcontroller through a serial interface. The 12-bit ADC does parallel sampling by using a method known as sample-and-hold.

C. Firmware and Software

We used a Teensy 4.0 to connect to the 4 audio channels. The Teensy is able to sample all 4 channels at a speed of 44.1kHz. To sample at a constant rate we used the timer interrupts on the Teensy. This allows us to know the sampling rate will not deviate. It then sends the 4 different values to a computer through serial and the computer does the HDR compositing so that we can record the original data and HDR data for offline processing for comparison. Alternatively, the HDR compositing was calculated directly on the Teensy for real-time performance benchmarks.

III. HDR Audio Algorithm

One method to compute HDR audio is a weighted average of scaled audio signals. The scale factor is determined based on the gain applied to the channel, and the maximum gain across all channels. If G represents the set of all gains, then the scale factor for a signal scaled by g_i is $\frac{max(G)}{g_i}$. This scales each signal so that relative to each other, they faithfully represent the true analog signal. The full HDR composite function is below.

$$HDR[n] = \frac{\sum_i \frac{c_i[f_i[n]]max(G)f_i[n]}{g_i}}{\sum_i c_i[f_i[n]]} \tag{2}$$

Here, c_i represents the certainty function. The certainty function represents the degree of certainty that is placed for each given signal measurement in a particular gain channel. It forms the basis for the HDR compositing, as the same signal measurement from different gain channels is combined into one. Due to the uncertainty present when saturation occurs, the certainty of the measurement is zero at this extreme. The certainty for higher gained unsaturated signal measurements should also be higher, since the gain quantization steps are smaller compared to low gained channels which would give a less accurate signal value.

IV. Results and Discussion

A. Importance of Gain Selection in Compositing

We examined the importance of selecting different input signal gains to the HDR compositing process. In this experiment, we limited the number of channels to two and varied the gains for the two channels from $1\times$ to $16\times$ in integer increments. These two-channel gain combinations were then used for the HDR signal reconstruction. To determine the quality of the reconstruction, the mean-squared errors (MSE) of each HDR composite relative to the reference audio signal were computed. These MSE values were then placed on a gain versus gain heat map, which we refer to as the gain map, to determine if an optimal gain combination exists.

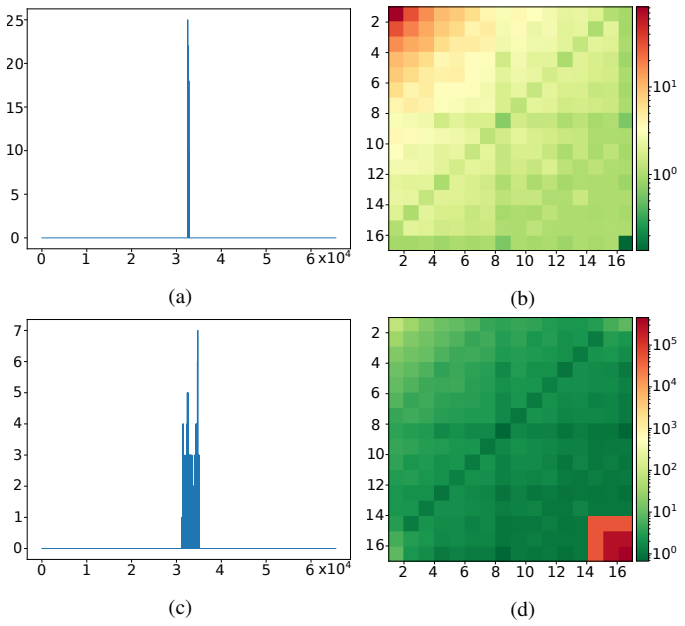


Fig. 6: Visualization of how the RDF affects the gain selection for HDR compositing. a) Histogram and b) MSE gain map for a quiet audio signal. c) Histogram and d) MSE gain map for a slightly louder audio signal.

The results indicate that the selection of gains is heavily dependent on the signal value distribution based on the ADC dynamic range. This has been referred to as the Range Density Function (RDF), described in [20], of the reference signal. Fig. 6 represents this finding. Fig. 6a is a histogram of a quiet audio input signal that uses only a narrow distribution of the ADC dynamic range. Using this as our source audio, we simulate lower 12-bit audio with different gains while accounting for ADC quantization and signal saturation. We use two simulated 12-bit audio signals to calculate the HDR composite signal and then calculate the MSE error. This is done for all combinations of gains from $1\times$ to $16\times$. Fig. 6b is the corresponding MSE gain map. Since the signal is quiet, the entirety of it can be captured using a high gain of $16\times$. The effect of having both channels with the same $16\times$ gain on the HDR compositing is similar to signal averaging. Since the signal is too quiet, the priority shifts from the need to extend the dynamic range of the ADC to the need of increasing the signal strength compared to noise.

We then use a louder audio signal distribution as shown in Fig. 6c. For this distribution, the optimal gains are $8\times$ and $16\times$. Fig. 6d shows the corresponding MSE gain map for the distribution. This shows that the importance of extending the dynamic range increases compared to the previous audio source. From the two examples, we can qualify that as the input signal value distribution expands, the optimal gain magnitudes become lower to cover a wider dynamic range. This shows that the need for HDR audio depends on the audio signal source. The need for dynamic gain adjustment on each individual gain channel could be useful in both increasing the dynamic range and the SNR of the audio signal.

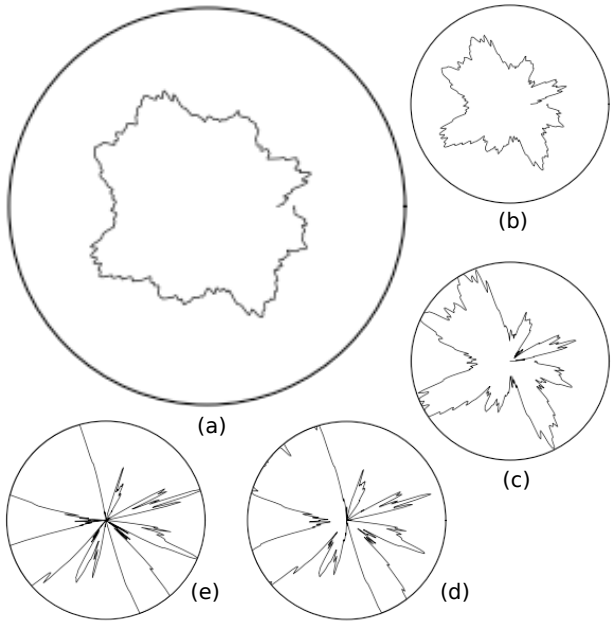


Fig. 7: SWIM Visualization of the optimal selected gains for a 4-channel system and its HDR Composition. a) HDR composite from the 4-channels. b) Audio gain of 2. c) Audio gain of 4. d) Audio gain of 8. e) Audio gain of 16.

Having explored the 2-channel case, a visual representation, by way of SWIM, of 4-channel HDR audio with optimally selected gains is shown in Fig 7. Fig 7a shows the HDR composite from the audio signals with gains of $2\times$, $4\times$, $8\times$ and $16\times$. The four 12-bit resolution signals combine to produce a 15-bit resolution HDR signal. As shown, the HDR signal does not contain any saturation and more details are captured due to the higher quantization resolution.

B. Influence of Number of Channels in Compositing

Intuitively, the more channels at different gain levels that are used for HDR compositing, the more accurate the information is obtained about a given signal. This implies that using more signals to generate an HDR signal results in a better representation of the true signal (i.e. a lower MSE). However, the extent of the improved representation with each addition of a signal channel in the compositing needs to be investigated. We consider whether a point exists where an increase in the number of channels does not provide any sufficient improvement to merit the extra monetary and computation cost.

Also, the number of required gains to best reconstruct a signal depends on the signal's range density function (RDF). If the RDF of the signal indicates a primarily low amplitude signal, then a single channel of audio at a high gain may be enough to perfectly reconstruct the signal. This is because high gain channels are able to accurately reconstruct low amplitude signals, and lower gains that are able to capture high amplitude signals are not required. However, if the RDF indicates a primarily high amplitude signal, then a single high gain signal will not be able to capture all points due to saturation. A number of differently gained channels are needed in this case, and increasing the number of channels will improve accuracy.

To confirm these observations experimentally, a program was created that recorded a 16-bit signal and reduced the resolution to 12-bits. Different gain levels were artificially applied to the reduced signal, and the resulting signals were combined to generate an HDR composite signal. To assess the quality of the signal reconstruction, the mean-squared error (MSE) was computed relative to the original 16-bit signal. This process was conducted with integer gain levels ranging from $1\times$ to $16\times$ and HDR compositions using 1 to 4 channels. In all cases, an HDR composition using four signal channels resulted in the lowest MSE.

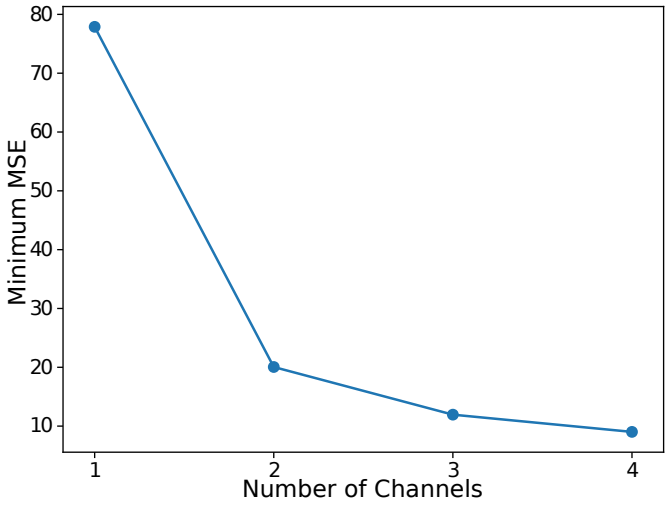


Fig. 8: Minimum MSE achieved using the optimal gain selection for multi-channel HDR composite.

Fig. 8 summarizes our observations using the optimal gains for HDR compositing. For each number of channels, all gain combinations were tested and the MSEs were collected. The minimum MSEs are shown in Fig. 8 for 1-, 2-, 3-, and 4-channel HDR composite. The MSE decreased with increasing number of channels however the degree of improvement is less at each increasing number. We limited the number of channels to 4-channels since the improvement between the 4- and 5-channel is most likely insignificant based on the improvement between the 3- and 4-channel. Thus, the decision to use 4-channel HDR signal compositing on the wearable HDR audio device.

C. Performance of Wearable HDR Audio in Real-time

1) Real-time performance: We used a Teensy 4.0 to determine the real-time performance of the wearable HDR audio device in Fig. 1 at the $44.1\,\mathrm{kHz}$ sampling rate. Table I summarizes the observed performance characteristics of the wearable HDR audio device.

TABLE I: Wearable HDR audio performance characteristics.

Metric	Value	Unit
Sampling Rate	44.1	kHz
Processing Time	18.4	µs
Times Real-time (TRT)	0.81	–
Number of Channels	4	–
Gain Settings	$1, 4, 16, 64$	–
HDR Resolution	18	bits
Hours of Operations	10	h

We obtained the processing time by timing each part of the code block that runs every time a sample is collected. This code block consists of three main parts: the sampling of multiple channels, the HDR calculation, and the transfer of data. The sampling of multiple channels were done through a serial interface with a clock rate of $4.8\,\mathrm{MHz}$. The total time taken to sample and retrieve 12-bit samples from 4 different channels was $13.4\,\mu\mathrm{s}$. The time required to calculate the HDR composite of 4 channels was on average $1\,\mu\mathrm{s}$ with the maximum time being $3\,\mu\mathrm{s}$. The HDR composite sample point was then transferred to another device for further processing. For this, we used serial communication to relay the data to a Raspberry Pi, which took a maximum of $2\,\mu\mathrm{s}$. The total processing time took a maximum of 18.4 microseconds to sample, average, and send the HDR sample to another device. Since the sampling frequency chosen was at $44.1\,\mathrm{kHz}$ ($22.6\,\mu\mathrm{s}$ between samples), the times real-time (TRT) value, which is defined as the ratio of the time taken to process the data to the actual duration of the data, would be about 0.81. This means that the device can run in real-time to output HDR audio.

Because the Teensy does not have a Bluetooth module, we use an external WiFi or Bluetooth chip to send the audio data wirelessly. An alternative would be to connect the Teensy to a Raspberry Pi which allows us to further process the data in a wearable form factor. The Raspberry Pi would allow the use of machine learning algorithms to generate speech-to-text data from the audio as well as save the data to the cloud which will be explored at a later time.

The HDR algorithm is not computationally demanding. Once we know the certainty function, we can create a lookup table with the certainty function. The certainty function determines the weighting of each channel for calculating a weighted average. The algorithm complexity is O(n) where n is the number of channels used. Because of this simplicity, the HDR algorithm can be done in a wearable device with little overhead. This gives an advantage over other signal reconstruction methods where the computation time is expensive.

2) Visualization using SWIM: To show the HDR audio performance, we recorded a steady $250\,\mathrm{Hz}$ tone at a $13\,\mathrm{kHz}$ sampling rate. To graph the figures shown in Fig. 9, we

made the plotted sample points equal to a multiple of the period of the sample tone. The period is 52 samples for a 250 Hz signal with a sampling rate of 13 kHz. This way, as the data is being plotted with no overlap, if the graphed sample points is a multiple of the period, the graph will be a sitting wave. As the physical distance between the microphone and the speaker moves, the graph will rotate either clockwise or counterclockwise depending on the direction of movement. In Fig. 9, the SWIM graphs show the phase shift as the distance from the speaker is incremented by 5in. This phase shift will be a function of distance and will be related to the signal's wavelength. Fig. 9a shows the HDR signal while the 8× gain signal channel is shown in Fig. 9b. The phase shift can be seen clearly in both scenarios, but in the HDR data, the signal is not saturated and contains more information about the signal. The single-channel gain saturates at these distances which leads to poorer audio quality.

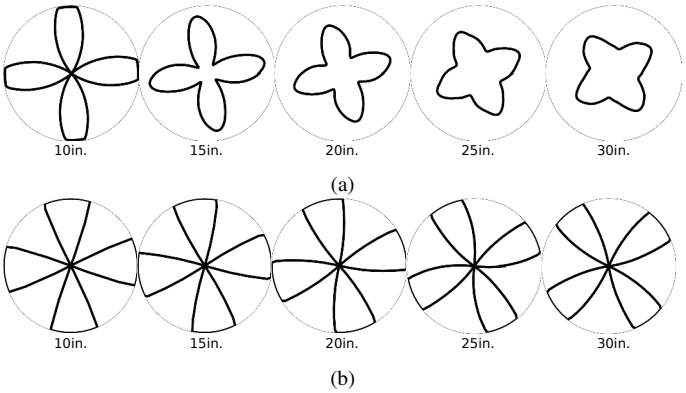


Fig. 9: Rotation in phase angle when the 250 Hz audio source is physically moved away from the microphone at 5in. increments. a) shows the HDR audio signal. b) shows a signal channel of 8× gain.

In Fig. 10, we overlayed the SWIM outputs to explicitly show the phase and amplitude change as the distance changes. The phase change, θ, is about 45° in the counter-clockwise direction as the distance between the audio transmitter and receiver increases from 5in. to 35in. Using the speed of sound at room temperature, the wavelength of a 250 Hz sound wave is about 1.4 m which is equivalent to 54in. The change in phase should then correspond to a difference in 27in. since 90° corresponds to a full period. The actual distance measured was about 30in. This 3in difference could be due to human error in measurements and error in the wavelength due to the environment. Thus, audio signals through the use of SWIM visualization could be used for distance measurement.

V. Conclusion

In this paper, we presented a novel wearable hardware device that increased the resolution of a multichannel ADC. This increased resolution effectively expanded its dynamic range. SWIM visualization was incorporated to better understand how HDR compositing improved audio performance and how audio can be used to measure distance. Additionally, we showed the importance of gain selection and the influence of the number of input channels in the HDR compositing.

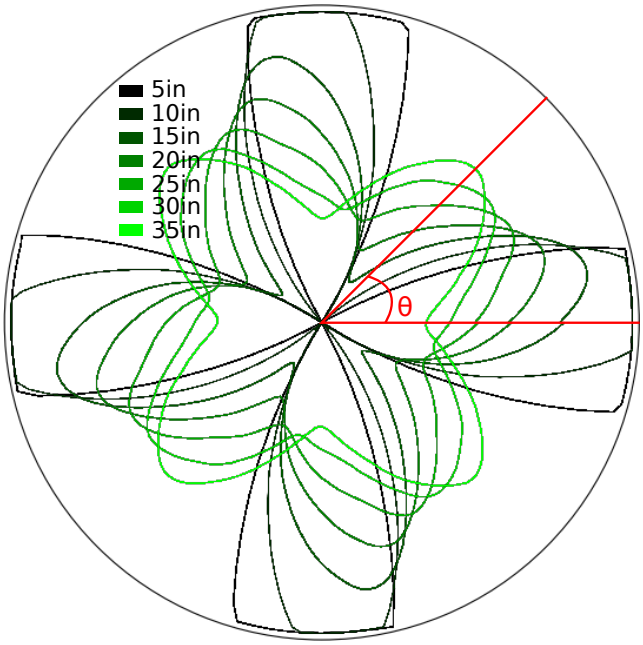


Fig. 10: Phase shift and amplitude change as the 250 Hz audio tone moves farther away from the microphone in 5in. increments. The patterns are the HDR audio composites at each distance. θ is a 45° phase shift by comparing the 5in. to the 35in. audio SWIM.

VI. Future Work

A. Speech, Tone and Emotion Recognition

Given a sample of recorded human speech, modern automatic speech and emotion recognition algorithms are able to detect and map features into the words, tone, and emotion of the speaker [31]–[33]. In the future, we will investigate the effect of increased resolution on the ability and performance to detect both speech and emotion in a sound sample across a wide range of volumes.

B. Assistive Technology Hearing Aid

Although the methods in this paper are able to effectively increase the dynamic range of recording devices, the spectral content of the sound being detected is not accounted for. Many people experience hearing loss as they become older. Each individual has a different auditory frequency response which has led to many developments in personalized hearing devices [34]–[37]. HDR audio can involve the ability to not only hone in on different volumes of a signal but also the relative volumes of frequencies present. In this way, we can create a personalized frequency response for every user for our device with the added benefit of a higher dynamic range.

C. Assistive Technology Auditory Memory Prosthetic

Our brain's capacity to remember daily experiences is not infinite. The use of wearable HDR audio devices for life-like audio recordings may provide a more immersive auditory memory prosthetic. The higher dynamic range audio, which can resemble professional audio equipment, would be able to capture background sounds (i.e. birds singing in the distance) that would not be otherwise captured. This enables the user

to recall their memory more easily with the addition of more audio cues in the recording [38], [39].

REFERENCES

[1] S. J. Godsill and P. J. W. Rayner, "Frequency-based interpolation of sampled signals with applications in audio restoration," in *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, 1993, pp. 209–212 vol.1.

[2] S. Nawab, T. Quatieri, and J. Lim, "Signal reconstruction from short-time fourier transform magnitude," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 4, pp. 986–998, 1983.

[3] F. Elvander, J. Swärd, and A. Jakobsson, "Grid-less estimation of saturated signals," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*, Oct 2017, p. 372–376.

[4] I. Stanković, M. Daković, and C. Ioana, "Time-frequency signal reconstruction of nonsparse audio signals," in *2017 22nd International Conference on Digital Signal Processing (DSP)*, Aug 2017, pp. 1–5.

[5] M. Harvilla and R. M. Stern, "Least squares signal declipping for robust speech recognition," in *INTERSPEECH*, 2014.

[6] Han Lin and S. Godsill, "The multi-channel AR model for real-time audio restoration," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2005.*, 2005, pp. 335–338.

[7] M. J. Harvilla and R. M. Stern, "Efficient audio declipping using regularized least squares," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 221–225.

[8] P. Smaragdis, "Dynamic range extension using interleaved gains," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 5, pp. 966–973, July 2009.

[9] L. Christodoulou, J. Lane, and T. Kasparis, "Dynamic range extension using multiple A/D converters," in *2010 4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, 2010, pp. 1–4.

[10] M. Granados, B. Ajdin *et al.*, "Optimal HDR reconstruction with linear digital cameras," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2010, pp. 215–222.

[11] J. Kuang, G. M. Johnson, and M. D. Fairchild, "icam06: A refined image appearance model for HDR image rendering," *Journal of Visual Communication and Image Representation*, vol. 18, no. 5, pp. 406–414, 2007.

[12] M. A. Ali, T. Ai *et al.*, "Comparametric HDR (high dynamic range) imaging for digital eye glass, wearable cameras, and sousveillance," in *2013 IEEE International Symposium on Technology and Society (ISTAS): Social Implications of Wearable Computing and Augmediated Reality in Everyday Life*. IEEE, 2013, pp. 107–114.

[13] M. A. Ali and S. Mann, "Comparametric image compositing: Computationally efficient high dynamic range imaging," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2012, pp. 913–916.

[14] J. Huang, V. Rampersad, and S. Mann, "High dynamic range tone mapping based on per-pixel exposure mapping," in *2013 IEEE International Symposium on Technology and Society (ISTAS): Social Implications of Wearable Computing and Augmediated Reality in Everyday Life*, June 2013, pp. 98–106.

[15] R. C. H. Lo, S. Mann *et al.*, "High dynamic range (HDR) video image processing for digital glass," in *Proceedings of the 20th ACM international conference on Multimedia*, 2012, pp. 1477–1480.

[16] S. Mann, R. C. H. Lo *et al.*, "Realtime HDR (high dynamic range) video for eyetap wearable computers, fpga-based seeing aids, and glasseyes (eyetaps)," in *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2012, pp. 1–6.

[17] T. Van Vo and C. Lee, "Robust HDR video synthesis using superpixel-based illumination invariant motion estimation," in *2018 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia)*, June 2018, pp. 206–212.

[18] H. Najaf-Zadeh, M. Budagavi, and E. Faramarzi, "VR+HDR: A system for view-dependent rendering of HDR video in virtual reality," in *2017 IEEE International Conference on Image Processing (ICIP)*, Sep. 2017, pp. 1032–1036.

[19] K. Myszkowski, R. Mantiuk, and G. Krawczyk, *High Dynamic Range Video*. Morgan & Claypool, 2008. [Online]. Available: https://ieeexplore.ieee.org/document/6812656

[20] R. Janzen and S. Mann, "High dynamic range simultaneous signal compositing, applied to audio," in *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, Apr 2012, p. 1–6.

[21] ——, "Extreme-dynamic-range sensing: Real-time adaptation to extreme signals," *IEEE MultiMedia*, vol. 24, no. 2, p. 30–42, Apr 2017.

[22] S. Mann, "Phenomenological augmented reality with the sequential wave imprinting machine (SWIM)," in *2018 IEEE Games, Entertainment, Media Conference (GEM)*. IEEE, 2018, pp. 1–9.

[23] P. Scourboutakos, M. H. Lu *et al.*, "Phenomenologically augmented reality with new wearable led sequential wave imprinting machines," in *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, 2017, pp. 751–755.

[24] S. Mann, "Phenomenal augmented reality: Advancing technology for the future of humanity," *IEEE Consumer Electronics*, pp. cover + 92–97, October 2015.

[25] ——, "Wavelets and" chirplets": Time–frequency" perspectives" with applications," in *Advances in machine vision: strategies and applications*. World Scientific, 1992, pp. 99–128.

[26] S. Mann, J. C. Havens *et al.*, "All reality: Values, taxonomy, and continuum, for virtual, augmented, extended/mixed (x), mediated (x, y), and multimediated reality/intelligence," *Presented at the AWE 2018*, 2018.

[27] S. Mann, T. Furness *et al.*, "All reality: Virtual, augmented, mixed (x), mediated (x, y), and multimediated reality," *arXiv preprint arXiv:1804.08386*, 2018.

[28] S. Mann, P. V. Do *et al.*, "Electrical engineering design with the subconscious mind," *2020 IEEE International Conference on Human-Machine Systems*, In Press, 2020.

[29] S. Mann, D. E. Garcia *et al.*, "Visualizing electric machines with sequential wave imprinting machine," *2020 Symposium on Virtual and Augmented Reality (SVR)*, In Press, 2020.

[30] "Low-cost, micropower, sc70/sot23-8, microphone preamplifiers with complete shutdown," MAX4465–MAX4469, Maxim Integrated, Jun. 2012, rev. 2. [Online]. Available: https://www.cuidevices.com/product/resource/cma-4544pf-w.pdf

[31] C. N. Moridis and A. A. Economides, "Affective learning: Empathetic agents with emotional facial and tone of voice expressions," *IEEE Transactions on Affective Computing*, vol. 3, no. 3, pp. 260–272, 2012.

[32] J. Umamaheswari and A. Akila, "An enhanced human speech emotion recognition using hybrid of prnn and knn," in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 2019, pp. 177–183.

[33] P. Tzirakis, J. Zhang, and B. W. Schuller, "End-to-end speech emotion recognition using deep neural networks," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 5089–5093.

[34] B. S. Wilson and M. F. Dorman, "Cochlear implants: current designs and future possibilities," *J Rehabil Res Dev*, vol. 45, no. 5, pp. 695–730, 2008.

[35] L. M. Friesen, R. V. Shannon *et al.*, "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," *The Journal of the Acoustical Society of America*, vol. 110, no. 2, pp. 1150–1163, 2001.

[36] S. Rosen, A. Faulkner, and L. Wilkinson, "Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants," *The Journal of the Acoustical Society of America*, vol. 106, no. 6, pp. 3629–3636, 1999.

[37] B. Widrow, "A microphone array for hearing aids," *IEEE Circuits and Systems Magazine*, vol. 1, no. 2, pp. 26–32, 2001.

[38] J. Hoisko, "Context triggered visual episodic memory prosthesis," in *Digest of Papers. Fourth International Symposium on Wearable Computers*, 2000, pp. 185–186.

[39] S. Vemuri and W. Bender, "Next-generation personal memory aids," *BT Technology Journal*, vol. 22, no. 4, pp. 125–138, 2004.