

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.,

Petitioner,

v.

JAWBONE INNOVATIONS, LLC,

Patent Owner.

Patent No. 8,467,543

Filing Date: March 27, 2003

Issue Date: June 18, 2013

Inventor: Gregory C. Burnett, Nicolas J. Petit,
Alexander M. Asseily, and Andrew E. Einaudi

Title: MICROPHONE AND VOICE ACTIVITY DETECTION (VAD)
CONFIGURATIONS FOR USE WITH COMMUNICATION SYSTEMS

**JAWBONE INNOVATIONS, LLC'S
PRELIMINARY RESPONSE**

Case No. IPR2023-01291

TABLE OF CONTENTS

	<u>Page(s)</u>
I. INTRODUCTION	1
II. THE '543 PATENT	1
III. THE RELEVANT ALLEGED PRIOR ART	5
A. Andrea (U.S. Patent No. 5,732,143)	5
B. Hussain (A New Metric for Selecting Sub-band Processing in Adaptive Speech Enhancement Systems)	7
C. Holzrichter (U.S. Patent No. 6,006,175)	10
IV. CLAIM CONSTRUCTION	11
V. LEVEL OF ORDINARY SKILL IN THE ART	11
VI. ARGUMENT	12
A. GROUND 1: Andrea, Hussain, and Holzrichter Do Not Render Obvious Claims 1-2 and 5-26	12
1. Andrea, Hussain, and Holzrichter Do Not Render Obvious a “a voice detection subsystem configured to receive voice activity signals that includes information associated with human voicing activity” where “the voice detection subsystem is configured to receive the voice activity signals using a sensor independent from the microphone array” [1A, 1I]	12
2. Andrea, Hussain, and Holzrichter Fail to Disclose or Render Obvious “the voice detection subsystem configured to automatically generate control signals using the voice activity signals;” [1B]	15

3.	Andrea, Hussain, and Holzrichter Fail to Disclose “components of the denoising subsystem automatically selecting at least one denoising method appropriate to data of at least one frequency subband of the acoustic signals using the control signals and processing the acoustic signals using the selected denoising method to generate denoised acoustic signals” 1[G]	17
4.	Andrea, Hussain, and Holzrichter Do Not Render Obvious “wherein the denoising method includes generating a noise waveform estimate associated with noise of the acoustic signals and subtracting the noise waveform estimate from the acoustic signal when the acoustic signal includes speech and noise,” [1H]	21
5.	Andrea, Hussain, and Holzrichter Fail to Disclose “wherein the voice detection subsystem is configured to receive the voice activity signals using a sensor independent from the microphone array and to output the control signals generated from the voice activity signals to the denoising system, the denoising system configured to use the control signals to denoise the acoustic signals from the microphone array” 1[I]	23
6.	Andrea, Hussain, and Holzrichter do not Disclose or Render Obvious Claim 26	28
B.	GROUND 2-3: These Grounds Fail Based on the Failure of Ground 1	29
VII.	CONCLUSION	29

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Chemours Co. FC, LLC v. Daikin Indus., Ltd.</i> , 4 F.4th 1370 (Fed. Cir. 2021)	26
<i>Intel. Corp. v. Qualcomm Inc.</i> , 21 F.4th 784 (Fed. Cir. 2021)	16
<i>PAR Pharm., Inc. v. TWI Pharms., Inc.</i> , 773 F.3d 1186 (Fed. Cir. 2014)	13
<i>PersonalWeb Techs., LLC v. Apple, Inc.</i> , 917 F.3d 1376	13

I. INTRODUCTION

On August 31, 2023, Meta Platforms, Inc. (“Meta” or “Petitioner”) filed a Petition requesting *inter partes* review of claims 1-26 (“Challenged Claims”) of U.S. Patent No. 8,467,543 (the “’543 Patent”) (Ex. 1001). Paper 1 (“Petition” or “Pet.”). The Declaration of David V. Anderson, Ph.D. (Ex. 1003) accompanied the Petition. On September 11, 2023, the Board issued a Notice of Filing Date Accorded for the Petition and set the time for filing patent owner’s preliminary response. Paper 3.

The Board should deny institution of the Petition because its proposed obviousness combinations fail to disclose numerous limitations relating to selecting a denoising method using control signals from a voice detection subsystem. In particular, Petitioner mischaracterizes both Holzrichter and Hussain to suggest that the voice activity detector mentioned in Hussain would be used to select the denoising method. In actuality, Hussain uses its MSC algorithm to perform the selection, and does not use the voice activity detector at all in the process. These deficiencies in the Petition are fatal to all Grounds.

II. THE ’543 PATENT

The ’543 Patent claims systems for performing acoustic noise suppression in a multiple microphone system with the use of a voice activity detector. ’543 Patent, Abstract (Ex. 1001). Embodiments of the system use two microphones and a voice activity detection (VAD) system, which includes both a VAD device and a VAD

algorithm. *Id.*, 5:45-50. The signals from the primary and secondary microphones are filtered into multiple frequency subbands, and each subband is denoised using its own adaptive filter. *Id.*, 16:31-36. The VAD system uses algorithms for processing the voice activity signals from the VAD device and outputs control signals used to adapt the noise suppression system to the received signals. *Id.*, 16:46-54, 19:53-64. The control signals can be used to determine when to adapt the transfer functions used in the denoising process to the presence or absence of speech. *Id.*, 14:51-52, 16:63-67.

The '543 Patent describes the embodiment of Figure 1:

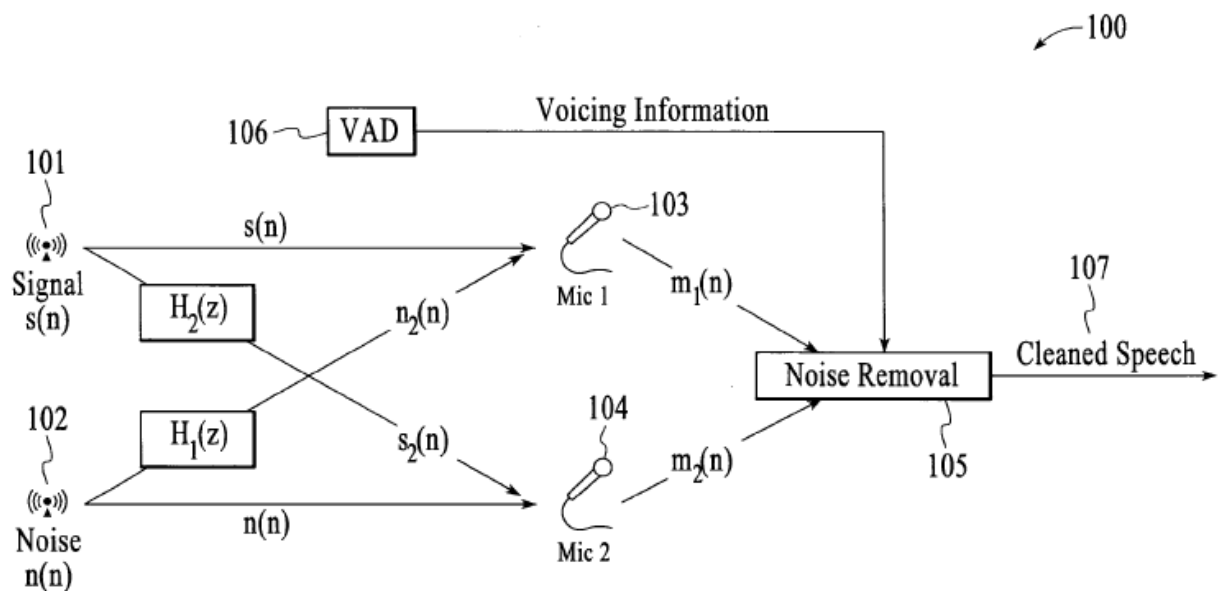


FIG.1

'543 Patent, FIG. 1.

In this embodiment, “[t]he signal processing system 105 includes two microphones MIC 1 103 and MIC 2 104 that receive signals or information from at least one speech source 101 and at least one noise source 102.” *Id.*, 14:43-46. “The path $s(n)$ from the speech source 101 to MIC 1 and the path $n(n)$ from the noise source 102 to MIC 2 are considered to be unity,” which means that MIC 1 is presumed to receive the entire speech signal and MIC 2 is presumed to receive the entire noise signal. *Id.*, 14:46-48. $H_1(z)$ is a transfer function that represents the amount of the noise signal received at MIC 1, and $H_2(z)$ is a transfer function that represents the amount of the speech signal that is received at MIC 2. *Id.*, 14:48-50.

The ’543 Patent teaches that “[t]he acoustic information coming into MIC 1 is denoted by $m_1(n)$. The information coming into MIC 2 is similarly labeled $m_2(n)$. In the z (digital frequency) domain, we can represent them as $M_1(z)$ and $M_2(z)$. Thus

$$“M_1(z)=S(z)+N(z)H_1(z)$$

$$M_2(z)=N(z)+S(z)H_2(z).”$$

Id., 14:52-60.

The ’543 Patent uses “[a] VAD signal 106, derived in some manner, . . . to control the method of noise removal.” *Id.*, 14:51-52. Using the VAD allows for simplification of these equations such that the equation for the sounds signal $S(z)$ can be derived as:

$$S(z) = \frac{M_1(z) - M_2(z)H_1(z)}{1 - H_2(z)H_1(z)}.$$

Id., 15:60-62.

Assuming the situation where “ $H_2(z)$ is quite small,” *i.e.*, that the amount of the speech signal reaching MIC 2 is very low, and assuming that “ $H_1(z)$ is less than unity,” *i.e.*, that less than the entire noise signal reaches MIC 1, the speech signal can be calculated as:

$$S(z) \approx M_1(z) - M_2(z)H_1(z).$$

Id., 16:2-3.

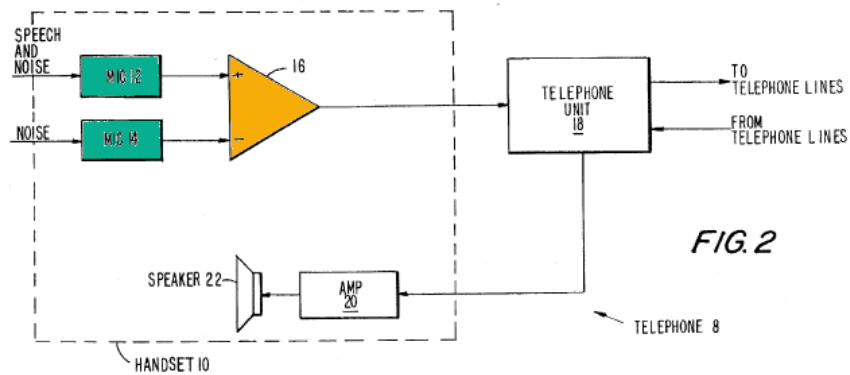
The '543 Patent teaches that its noise suppression system is an improvement over the prior art because it uses an accurate voice activity detection signal in the denoising process. *Id.*, 2:7-12. The '543 Patent teaches that it is crucial to implement “a sufficiently accurate VAD . . . to adapt the coefficients of H_1 (noise only) and H_2 (if needed, when speech is being produced).” *Id.*, 16:55-67. If the VAD is inaccurate and speech is occurring during the calculation of the coefficients of H_1 (noise only), it “would result in the removal of a large part of the speech energy from the signal of interest,” resulting in “signal distortion and reduction (de-signaling).” *Id.*

Using multiple subbands in processing the acoustic signals results in significant noise suppression. *Id.*, 16:9-11. The signals from MIC 1 and MIC 2 are filtered into multiple subbands and the resulting data from each subband is sent to its own adaptive filter. *Id.*, 16:31-36. The results from each subband are added back together to form the final denoised signal. *Id.*, 16:38-40.

III. THE RELEVANT ALLEGED PRIOR ART

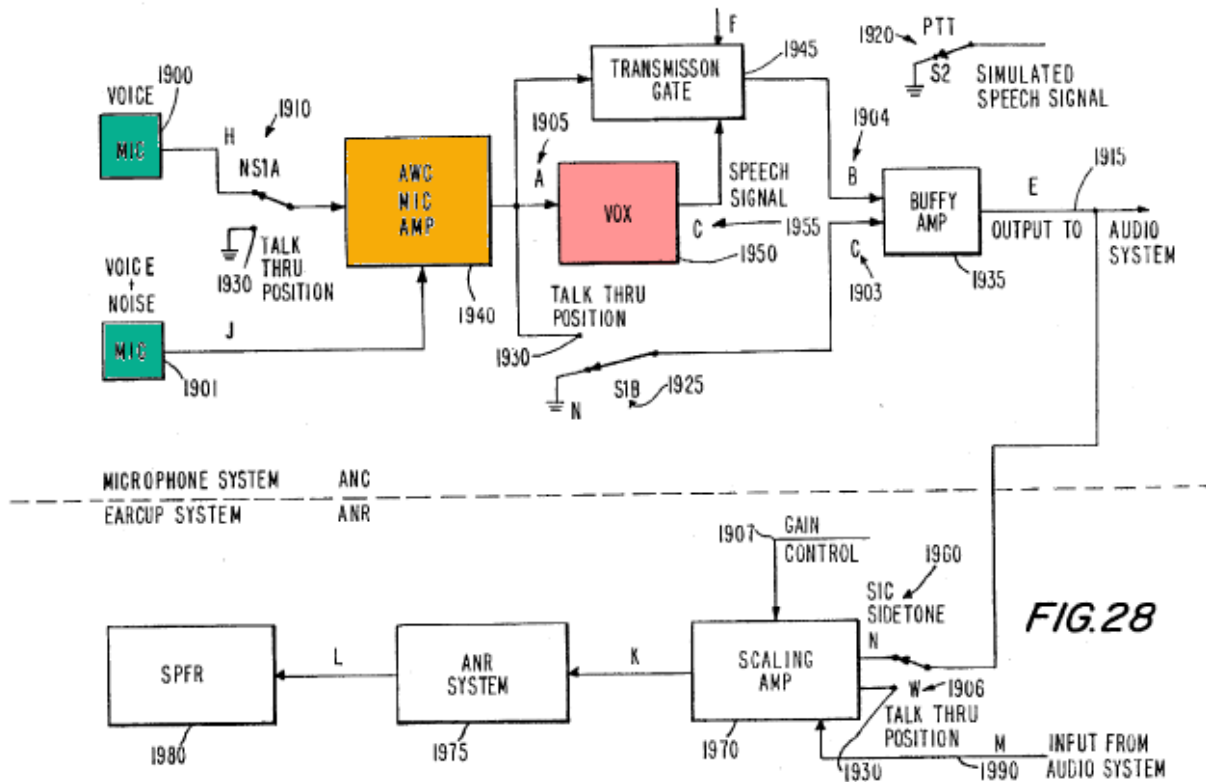
A. Andrea (U.S. Patent No. 5,732,143)

U.S. Patent No. 5,732,143 (Ex. 1005) (“Andrea”) is directed to a primitive two-microphone handset or headset apparatus. Ex. 1005, FIG. 2. The first microphone receives speech and noise, and the second microphone receives only noise. *Id.*, 12:55-62. The speech-and-noise signal is denoised in the op-amp by subtracting the noise signal of the second microphone from the speech and noise signal of the first microphone. *Id.*, 12:63-13:1. This architecture is shown in Andrea Fig. 2, reproduced below with microphones highlighted in green and op-amp in yellow.



Id., FIG. 2 (annotations added).

Petitioner further relies on a particular embodiment of Andrea, illustrated in Fig. 28, that includes a VOX circuit. Pet., 6 (“A ‘VOX circuit’ senses speech in a ‘headset having a ‘talk-thru’ capability.’” (quoting Ex. 1005, 33:44-35:7). Fig. 28 is reproduced below, with the VOX circuit highlighted in red:



Ex. 1005, FIG. 28 (annotations added).

B. Hussain (A New Metric for Selecting Sub-band Processing in Adaptive Speech Enhancement Systems)

The paper *A New Metric for Selecting Sub-band Processing in Adaptive Speech Enhancement Systems* (“Hussain”) (Ex. 1006) was purportedly published in 1997. Hussain is directed to “A multi-microphone adaptive speech enhancement system employing diverse sub-band processing.” Hussain, 2611.

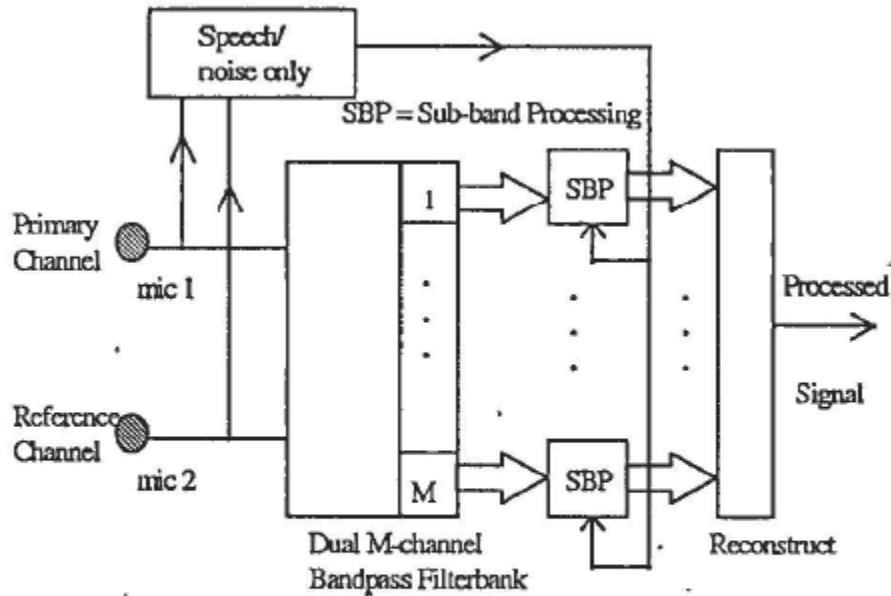


Figure 1: The MMSBA system

Hussain discloses that it has developed a “new robust metric . . . which is capable of real-time implementation, in order *to automatically select* the best form of processing within each sub-band.” *Id.* (emphasis added). “In order to realize this, a robust practical metric has been developed, which is capable of realtime implementation, based on the inter-channel Magnitude Squared Coherence (MSC) relationship *in order to automatically select the best SBP [Sub-Band Processing] option.*” *Id.* (emphasis added). Based on the MSC, Hussain automatically chooses one of three options for subband processing: “(i) no processing; (ii) intermittent coherent noise canceller; and (iii) continuous incoherent noise canceller.” *Id.*

Hussain “propose[s] the use of a modified MSC as a part of a system for *selecting the best SBP option* in a [multi-microphone sub-band adaptive] speech enhancement system.” Hussain, 2612 (emphasis added). Hussain derives the following equation for the MSC:

$$\overline{\text{MSC}}(f_k) = \frac{1}{I} \sum_{i=1}^I \text{MSC}(f_k, i)$$

Hussain, 2613. According to Hussain:

The above adaptively averaged MSC criterion can thus be used as an effective means for determining the level of correlation between the disturbing noises at various frequencies bands (by averaging the above MSC over each respective sub-band), during the noise alone period in intermittent speech. *The subsequent form of processing in each respective frequency band can therefore be selected on the basis of the inter-channel correlation.*

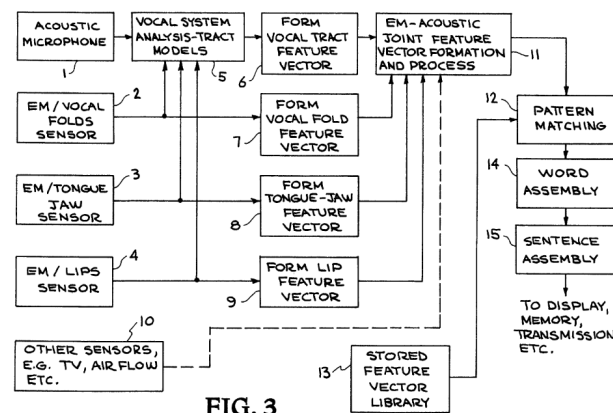
Id. (emphasis added).

Hussain uses the MSC to choose between using an “intermittent coherent noise canceller” and an “incoherent noise canceller.” Hussain, 2612-13. According to Hussain, “[o]n initial trials, a threshold value of 0.6 for the adaptive MSC has been found to be suitable for distinguishing between highly correlated and weakly correlated sub-band noise signals.” Hussain, 2613. Hussain teaches that when “the

noise power is significant and the noise between two channels is significantly correlated in a sub-band, then perform adaptive intermittent noise cancellation.” Hussain, 2612. But, “[i]f the noise power is significant but not highly correlated between the two channels in a sub-band, then the incoherent noise cancellation approach . . . may be applied during the noisy speech period.” *Id.*

C. Holzrichter (U.S. Patent No. 6,006,175)

U.S. Patent No. 6,006,175 (“Holzrichter”) (Ex. 1007) was filed on February 6, 1996, and issued on December 21, 1999. Holzrichter “relates generally to speech recognition and more particularly to the use of nonacoustic information in combination with acoustic information for speech recognition.” Ex. 1007, 1:12-15.



Holzrichter, FIG. 3.

Holzrichter discloses using EM sensors for measuring the positions and movements of certain organ positions, including the vocal folds, velum, tongue, and lips. Ex. 1007, 16:42-49. Feature vectors are generated which describe the signals

generated by the EM sensors which are then combined with feature vectors describing acoustic information “to produce a more complex, more information-rich feature vector describing the speech unit time frame.” *Id.*

Holzrichter’s pattern matches the combined feature vector against a feature vector library to identify the sound represented by the feature vector. Ex. 1007, 17:32-33. Words and sentences are assembled from the collection of identified sounds to yield the results of Holzrichter’s voice recognition system. *Id.*, 17:34-36.

Holzrichter teaches that the algorithm for determining when speech begins, i.e., the onset of speech, requires the use of an EM sensor and an acoustic sensor. Ex. 1007, 31:33-32:54. “The algorithm detects the onset of speech by using both an acoustic and one or more EM sensor modules at the same time.” *Id.*, 31:51-53. Although the EM sensors are effective for detecting voiced speech, the acoustic sensor is used to detect unvoiced speech when no vocal fold motion is detected. *Id.*, 32:7-10, Fig. 19, element 79.

IV. CLAIM CONSTRUCTION

For the purposes of this Preliminary Response only, Patent Owner believes that claim construction is not required to resolve any issues.

V. LEVEL OF ORDINARY SKILL IN THE ART

For the purposes of this Preliminary Response only, Patent Owner utilizes Petitioner’s proposed level of skill in the art—“a minimum of a bachelor’s degree in

computer engineering, computer science, electrical engineering, mechanical engineering, or a similar field, and approximately three years of industry or academic experience in a field related to acoustics, speech recognition, speech detection, or signal processing. Work experience can substitute for formal education and additional formal education can substitute for work experience.” Pet. at 4-5 (internal citations omitted).

VI. ARGUMENT

A. GROUND 1: Andrea, Hussain, and Holzrichter Do Not Render Obvious Claims 1-2 and 5-26

1. Andrea, Hussain, and Holzrichter Do Not Render Obvious a “a voice detection subsystem configured to receive voice activity signals that includes information associated with human voicing activity” where “the voice detection subsystem is configured to receive the voice activity signals using a sensor independent from the microphone array” [1A, 1I]

Petitioner’s combination fails to render obvious “a voice detection subsystem configured to receive voice activity signals that includes information associated with human voicing activity” [1A] because it relies on the microphone array instead of “a sensor independent from the microphone array.” See [1I]. Petitioner argues that Andrea alone or Andrea with Hussain meets [1A]’s limitation. Pet., 13.

At the outset, with respect to its Andrea alone argument, the Petition argues that a POSITA would have understood that an analysis (“*determining* whether the

microphone signal is useable audio”) is somehow a reception of signals (“receiving voice activity signals that include information associated with human voicing activity”). Pet., 13 (emphasis added). In particular, the Petition does not identify: i) what the voice activity signals are; and ii) how they are received. Moreover, the Petition does not explain how “determining” something is equivalent to “receiving.” To the extent the Petition is meant to argue that “determining whether the microphone signal is useable audio” *inherently* comprises “receiving voice activity signals that include information associated with human voicing activity,” it fails to set forth any argument or evidence that “the natural result flowing from the operation as taught would result in the performance of the [receiving] function.” *PersonalWeb Techs., LLC v. Apple, Inc.*, 917 F.3d 1376, 1382 (Fed. Cir. 2019 (quoting *PAR Pharm., Inc. v. TWI Pharms., Inc.*, 773 F.3d 1186, 1195 (Fed. Cir. 2014))).

Petitioner’s expert provides no further assistance. Dr. Anderson’s Declaration merely repeats the Petition verbatim, and provides no explanation or citation for his conclusory opinion. See Ex. 1003, ¶ 74.

The only mention of “signals” in Petitioner’s argument is with respect to the “signal at the output of the microphone.” Pet., 13; Ex. 1003, ¶ 74. But if the microphone signal is the voice activity signal, then it cannot meet limitation [11]’s requirement that “the voice detection subsystem is configured to *receive the voice activity signals using a sensor independent from the microphone array.*” And

Andrea contains no other sensors. Thus, Andrea cannot render this limitation obvious.

Petitioner alternatively argues that Andrea, with Hussain, renders this limitation obvious. Pet., 13-17. According to Petitioner, Hussain's statement that it "assumes" an effective VAD is available amounts to "receiving voice activity signals." *Id.*, 14 ("A POSITA would have understood detecting a 'speech plus noise period' corresponds to receiving voice activity signals that include information associated with human voicing activity."). As with its Andrea theory, Petitioner does not identify what it contends the voice activity signals are, nor does it describe the "sensor independent from the microphone array." And Petitioner does not attempt to show that Hussain inherently discloses such a sensor.

While Petitioner's identified combination for limitation [1A] does not include any configuration that includes a "sensor independent from the microphone array," it later attempts to backfill the limitation in its discussion of [1I]. That argument is discussed further below.

Accordingly, the Petition is deficient and the Board should decline to institute.

2. Andrea, Hussain, and Holzrichter Fail to Disclose or Render Obvious “the voice detection subsystem configured to automatically generate control signals using the voice activity signals;” [1B]

Petitioner’s combination fails to disclose or render obvious this limitation because it does not identify any control signals that are “automatically generate[d] using the voice activity signals.”

Petitioner first argues that Andrea, standing alone, renders this limitation obvious. Pet., 17-18. But, as discussed above, Andrea alone cannot be the voice detection subsystem because it does not receive voice activity signals using a sensor independent from the microphone array.

Petitioner next argues that Holzrichter and Hussain render this limitation obvious. Pet., 18-23. According to Petitioner, a POSITA would have used Holzrichter to implement the VAD whose existence Hussain assumes. *Id.*, 18. Holzrichter uses its EM sensor, along with acoustic sensors, to output a “feature vector.” Holzrichter, 17:24-31. As Holzrichter explains, a feature vector is the “numerical attributes of a set of . . . phonetic descriptions.” *Id.*, 1:24-25. The phonetic descriptions are attributes of a particular sounded phoneme, such as “formants, pitch, voiced/unvoiced, energy, nasality, and frication.” *Id.* 1:22-24.

Holzrichter’s system outputs the feature vector of a particular “sound unit, which are often defined as syllables, phonemes, or phone like units (PLU).” *Id.*,

10:2-8; 17:4-23. The feature vectors are compared to a library of feature vectors to identify sounds, syllables, phonemes, or phone-like units. *Id.*, 9:63-10:8, 17:27-33. “[W]ord assembly 14 and sentence assembly 15 can be performed and tested by word spellers, grammar and syntax correcting systems, and sent to where the user wishes.” *Id.*, 17:34-36.

The Holzrichter feature vector thus is not a “control signal” because it does not control anything. It is instead a series of numbers that are compared with other numbers. Neither the Petition nor Dr. Anderson explain why a “POSITA would have understood that automatically generating or forming the feature vector corresponding to the vocal folds signals is automatically generating control signals using the voice activity signals.” Pet., 20; Ex. 1003, ¶ 89. Instead, they simply assert their conclusion. A bald assertion would not be substantial evidence. *Intel. Corp. v. Qualcomm Inc.*, 21 F.4th 784, 797 (Fed. Cir. 2021) (stating that “[c]ertainly, conclusory expert testimony does not qualify as substantial evidence”) (citations and alterations omitted).

Accordingly, the Petition is deficient because the combination does not disclose or render obvious anything that “automatically generate[s] control signals using the voice activity signals.”

**3. Andrea, Hussain, and Holzrichter Fail to Disclose “components of the denoising subsystem automatically selecting at least one denoising method appropriate to data of at least one frequency subband of the acoustic signals using the control signals and processing the acoustic signals using the selected denoising method to generate denoised acoustic signals”
1[G]**

Petitioner’s combination does not disclose or render obvious “components of the denoising subsystem automatically selecting at least one denoising method appropriate to data of at least one frequency subband of the acoustic signals using the control signals.” The Petition relies on Hussain to meet this limitation, but Petitioner presents no evidence that Hussain discloses control signals generated from voice activity signals are used to automatically select a denoising method.

Claim 1 specifies that the “control signals” of this limitation are “automatically generated” by the “voice detection subsystem.” ’543 Patent, 24:3-7. Petitioner argues that the Holzrichter’s EM sensor is the Hussain VAD, and the feature vector is the control signal. Pet., 37-38.

Petitioner relies on Hussain to teach “automatically selecting at least one denoising method appropriate to data of at least one frequency subband of the acoustic signals using the control signals.” *Id.*, 35-38. Petitioner argues:

Hussain discloses the claimed automatically selecting at least one denoising method appropriate to data of at least one

frequency subband of the acoustic signals because it describes “automatically select[ing] the best form of processing within each [frequency] sub-band” to process the data in the sub-band. Ex. 1006, 10; Ex. 1003, ¶ 125.

Hussain discloses ***processing*** the acoustic signals of each sub-band using the selected denoising method to generate denoised acoustic signals by ***“perform[ing] adaptive intermittent noise cancellation” using a VAD and detecting “noise alone” and “speech plus noise” periods, as discussed in [1C]***. Ex. 1006, 11; *supra* § V.A.4.d; Ex. 1003, 126. “[D]uring the noise alone period” (when voicing activity is not detected), Hussain uses “an adaptive filter” to determine a “differential acoustic-path transfer function” for each sub-band. Ex. 1006, 11; Ex. 1003, ¶ 126. Then, “during the speech plus noise period” (when voicing activity is detected), Hussain uses the transfer function “to process the noisy speech signal” and generate and output “a noise reduced, filtered version of the sub-band speech signal” (denoised speech signal). Ex. 1006, 11; Ex. 1003, ¶ 126. Hussain then “[r]econstruct[s]” the denoised acoustic signal by combining each denoised sub-band signal into a single “Processed Signal.” Ex. 1006, 11, 13, Fig. 1; Ex. 1003, ¶ 126.

Id., 36-37 (emphasis added). According to Petitioner, “Hussain strongly suggests this process is performed using the control signals as recited in [1G], which Holzrichter confirms and renders obvious.” *Id.*, 37.

Contrary to the Petition, Hussain does not use the VAD to select a denoising method. As discussed above, Hussain teaches “the use *of a modified MSC* as a part of a system for *selecting the best SBP option* in a [multi-microphone sub-band adaptive] speech enhancement system.” Hussain, 2612 (emphasis added). Hussain characterizes this modified MSC as “a robust practical metric,” *id.*, 2611, which is derived using the following equation to “determine[e] the level of correlation between the disturbing noises at various frequencies bands (by averaging the above MSC over each respective sub-band)”:

$$\overline{\text{MSC}}(f_k) = \frac{1}{l} \sum_{i=1}^l \text{MSC}(f_k, i)$$

Id., 2612-13.

Hussain teaches that the “control signal” that determines which denoising method to use is the MSC. Hussain uses the MSC to choose between using an “intermittent coherent noise canceller” and an “incoherent noise canceller.” *Id.*, 2612-13. Hussain discloses that “[o]n initial trials, a threshold value of 0.6 for the adaptive MSC has been found to be suitable for distinguishing between highly correlated and weakly correlated sub-band noise signals.” *Id.*, 2613. According to Hussain, an intermittent coherent noise canceller is used when the sub-band noise signals are highly correlated. Hussain, 2612 (stating that the intermittent coherent

noise canceller is used “[i]f the noise power is significant and the noise between the two channels is *significantly correlated* in a sub-band.”) (emphasis added). On the other hand, an incoherent noise canceller is used when noise power is weakly correlated between two channels in a sub-band. Hussain, 2612 (stating that the incoherent noise canceller is used “[i]f the noise power is significant but *not highly correlated* between the two channels in a sub-band.”) (emphasis added). As such, Hussain expressly teaches that the MSC is used to select between the denoising methods disclosed in Hussain.

With a proper understanding of Hussain, the Petition is deficient because it identifies no evidence that Hussain uses control signals from a VAD to select a denoising method. While the Petitioner is correct that: i) a denoising method is selected and ii) that a VAD is present in the system, the Petition—and Hussain—contain no evidence that the VAD is used in the selection process. That is because Hussain exclusively uses its MSC to select the denoising method. Hussain, 11 (“A metric for selecting the appropriate SBP option is now derived next.”). The very next section is “2.1 The adaptively estimated Magnitude Squared Coherence (MSC) Metric.” *Id.*

Hussain only uses the VAD to determine whether to cancel noise or to adapt the filter during the intermittent noise cancellation SBP option. *Id.* It uses MSC to *select whether* to use intermittent noise cancellation or incoherent noise

cancellation. *Id.* The Petition tacitly admits as much by stating that Hussain “discloses *processing* the acoustic signals of each sub-band *using the selected denoising method* to generate denoised acoustic signals by ‘*perform[ing] adaptive intermittent noise cancellation*’ *using a VAD* and detecting ‘noise alone’ and ‘speech plus noise’ periods, as discussed in [1C].” Pet., 37. In other words, the Petition does not even assert that the VAD is used to *select* the method, only to process or perform the *already-selected* adaptive intermittent noise cancellation method.

Accordingly, the Petition is deficient because it presents no evidence that Hussain uses a VAD signal as a control signal to select a denoising method. Therefore, the Board should deny institution.

4. Andrea, Hussain, and Holzrichter Do Not Render Obvious “wherein the denoising method includes generating a noise waveform estimate associated with noise of the acoustic signals and subtracting the noise waveform estimate from the acoustic signal when the acoustic signal includes speech and noise,” [1H]

Ground 1 is deficient also because the combination of Andrea, Hussain, and Holzrichter does not render obvious generating a noise waveform estimate associated with noise of the acoustic signals and subtracting the noise waveform estimate from the acoustic signal.

The Petitioner argues that Andrea discloses generating a waveform estimate associated with noise when it “convert[s] the background noise supplied to the second microphone into a corresponding electric signal.” Pet. at 40. A portion of the specification cited by Petitioner states that “[t]he background noise is supplied to the second microphone 14 and converted therein into a corresponding electrical signal which is thereafter supplied to the minus terminal of the op-amp 16.” Andrea, 12:59-62. But this disclosure merely describes the typical operation of a microphone: converting acoustic signals to electric signals. It does not disclose generating an estimate of a noise waveform.

Petitioner’s argument would only have merit if claim term [1H], as denoted by Petitioner, were properly construed as “wherein the denoising method includes *subtracting the signal received at the noise microphone* from the acoustic signal when the acoustic signal includes speech and noise.” Petitioner presents no argument or evidence that “generating a noise waveform estimate” should be construed as nothing more than using a noise microphone. Petitioner’s position is particularly tenuous in light of the ’543 Patent’s use of transfer functions to estimate the noise waveform. *See* ’543 Patent, 14:40-16:3.

Petitioner also argues, through its expert, that the operation of weighting the amplitude of the signal from the noise microphone to match the amplitude of the output of the first microphone generates a noise waveform estimate. Pet at 40 (citing

Andrea, 15:29-34 and Ex. 1003, ¶ 133). Again, Petitioner presents no evidence or reasoning why simply matching the amplitude of two signals comprises “generating a noise waveform estimate.” Simply making the noise signal louder so that the subtraction operation will be more useful does not seek to estimate a noise waveform.

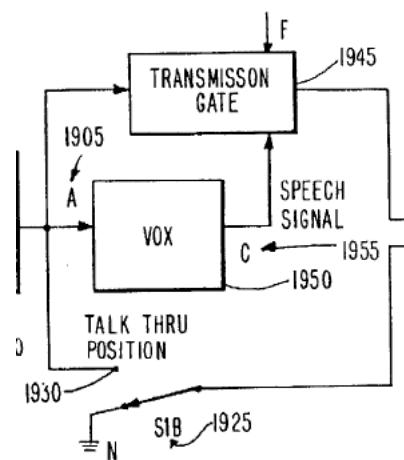
Accordingly, Ground 1 is deficient because the combination of Andrea, Hussain, and Holzrichter does not render obvious generating a noise waveform estimate associated with noise of the acoustic signals and subtracting the noise waveform estimate from the acoustic signal.

5. Andrea, Hussain, and Holzrichter Fail to Disclose “wherein the voice detection subsystem is configured to receive the voice activity signals using a sensor independent from the microphone array and to output the control signals generated from the voice activity signals to the denoising system, the denoising system configured to use the control signals to denoise the acoustic signals from the microphone array” 1[I]

The Petition is further deficient because Petitioner’s combination does not disclose or render obvious “wherein the voice detection subsystem is configured to receive the voice activity signals using a sensor independent from the microphone array and to output the control signals generated from the voice activity signals to

the denoising system.” Petitioner argues that Andrea alone or Andrea with Hussain and Holzrichter render this limitation obvious.

As discussed above, Andrea alone cannot render this limitation obvious because Andrea uses the microphone array and not a separate sensor. Moreover, Andrea’s VOX circuit outputs only a speech audio signal and not any control signal.



Andrea, FIG. 2 (excerpted).

Petitioner asserts that Holzrichter renders obvious a voice detection subsystem using a sensor independent from the microphone array to receive the voice activity signals and output the control signals. Pet. at 41. However, a POSITA would not be motivated to remove Holzrichter’s inventive functionality to only extract its EM sensor.

In a hindsight attempt to meet this limitation using Holzrichter, Petitioner disassembles Holzrichter’s sensor bank to identify a single sensor, the vocal folds EM sensor 23, as the “sensor independent from the microphone array.” Pet. at 41-

43. Specifically, Petitioner argues that the claimed voice detection subsystem includes the EM sensor 23 and the functional block that forms the vocal fold feature vector from the output of the EM sensor 23. Pet 43 (citing Ex. 1007, FIG. 3, elements 2 and 7). The EM sensor 23 detects motion in the vocal folds and glottis. Ex. 1007, 17:62-63. Petitioner argues that the EM sensor 23 is independent from the microphone array because Holzrichter discloses a microphone 24 which obtains simultaneous acoustic information but which is shown separately in Figure 4 of Holzrichter. Pet. at 42.

Holzrichter's inventive concept is to combine feature vectors obtained from the EM sensor with those obtained from acoustic microphones, and thus obtain an accurate description of the recorded sound unit to be used in speech recognition. *E.g.*, Holzrichter, Abstract (“***By simultaneously recording EM wave reflections and acoustic speech information***, the positions and velocities of the speech organs as speech is articulated can be defined for each acoustic speech unit.”); 7:21-24 (“Accordingly it is an object of the invention to provide method and apparatus for speech recognition ***using nonacoustic information in combination with acoustic information.***”). Holzrichter explains that both the EM sensor 23 and the accompanying acoustic sensor are both necessary to accurately detect speech.

Holzrichter's acoustic sensor is necessary because it detects unvoiced speech that is missed by the EM sensor monitoring vocal-fold motion. “[t]he issue for this

algorithm is that some words begin with unvoiced sounds such as ‘s’ in ‘sam’ or in ‘sixteen’; thus the algorithm must be able to back up *to catch sounds that may be missed by the vocal fold motion onset detection by the EM sensor.*” Ex. 1007, 31:53-57 (emphasis added). The algorithm recognizes that unvoiced speech is present when the acoustics detect speech when vocal-fold motion is absent:

If acoustics and vocal fold motion are present, then a voiced speech unit has occurred and the unvoiced timer, T, is set to zero in box 78, and processing on frame t_i continues. If satisfactory acoustics, *but no vocal fold motion occurs* in box 72, the sound in frame t_i is labeled as nonvoiced speech in box 79 and processing continues.

Id., 32:4-10 (emphasis added). Despite this disclosure, the Petition argues that a POSITA would include **only** the EM sensor from Holzrichter and use that, alone, as a voice activity detector, while leaving the acoustic sensor behind. A POSITA, considering Andrea, Hussain, and Holzrichter, would not be motivated to remove Holzrichter’s acoustic functionality and thus defeat Holzrichter’s inventive concept. *Chemours Co. FC, LLC v. Daikin Indus., Ltd.*, 4 F.4th 1370, 1376 (Fed. Cir. 2021) (reversing Board judgment of obviousness where proposed combination “would necessarily involve altering the inventive concept” of the prior art reference). Indeed, the only possible motivation for importing a crippled Holzrichter into Andrea/Hussain is a hindsight-based desire to meet the claims.

Moreover, it would not have been obvious to use Holzrichter's control signals from the EM sensor alone in Hussain's sub-band processing noise-cancellation system because those control signals would not have accurately identified periods of speech and no speech as required by Hussain. While Hussain does not identify a specific device to use as a voice activity detector, his system includes a "speech/noise only" block whose output identifies two time periods: a "speech plus noise period" and a "noise alone period." Ex. 1006 at 11, 13, FIG. 1. The output from this block is used by Hussain to create the adaptive filters used to denoise the noisy speech signal. *Id.* at 11.

Holzrichter's EM sensor without an accompanying acoustic sensor would detect only voiced speech and would not provide an accurate "noise only" signal to Hussain because periods of unvoiced speech would be undetected as speech and mischaracterized as noise.

A POSITA would not be motivated to combine Holzrichter's control signals from the EM sensor with Hussain because it would reduce the effectiveness of Hussain's denoising system. The purpose of the "effective voice activity detector (VAD)" in Hussain is to identify the noise only period that is used to determine the adaptive filters that model the transfer function between the microphones during the noise only period. *Id.* at 11. If the VAD used by Hussain is inaccurate and speech is occurring when Holzrichter's control signals indicate a noise only period, the

adaptive filters will not properly model the transfer function and, therefore, the noise suppression of the Hussain system would suffer.

Even if a POSITA were somehow motivated to combine the control signals from Holzrichter's EM sensor with the Hussain system, a POSITA would not have reasonably expected success in the combination because it would not have yielded an effective denoising system for the reasons discussed above: a sensor which cannot detect unvoiced speech will not properly identify periods of no speech as required by Hussain. A POSITA reviewing the Holzrichter reference would recognize that to detect unvoiced speech, the output of the EM sensor was paired with the output of an acoustic sensor, but Petitioner's combination explicitly excludes the acoustic sensor and its output from its identification of the claimed voice detection system.

Accordingly, Petitioner's combination does not disclose or render obvious this limitation.

6. Andrea, Hussain, and Holzrichter Do Not Disclose or Render Obvious Claim 26

Petitioner relies on the same arguments for limitations [26A]-[26I] as it does for claim 1. For the same reasons as stated above, Petitioner's combination therefore does not render obvious claim 26.

B. GROUND 2-3: These Grounds Fail Based on the Failure of Ground 1

Petitioner relies on the combination of Andrea, Hussain, and Holzrichter with Stevens (Ground 2) or Alcivar (Ground 3). Neither Grounds 2 nor 3 remedies the deficiencies of Ground 1. Accordingly, Grounds 2 and 3 are deficient for the same reasons discussed with respect to Ground 1.

VII. CONCLUSION

As discussed above, Petitioner's combinations fail to disclose key limitations of each claim of the '543 Patent, and therefore, do not render obvious any claims. The Board should, therefore, deny the Petition.

Respectfully submitted,

Dated: December 11, 2023 By: /Peter Lambrianakos /
Peter Lambrianakos (Reg. No. 58,279)
FABRICANT LLP
411 Theodore Fremd Avenue,
Suite 206 South
Rye, New York 10580
Tel. 212-257-5797
Fax. 212-257-5796
Email: plambrianakos@fabricantllp.com

CERTIFICATE OF WORD COUNT

The undersigned hereby certifies that the portions of the above-captioned JAWBONE INNOVATIONS, LLC'S PRELIMINARY RESPONSE specified in 37 C.F.R. § 42.24 has 5,289 words in compliance with the 14,000 word limit set forth in 37 C.F.R. § 42.24. This word count was prepared using Microsoft Word for Office 365.

December 11, 2023

By: /Peter Lambrianakos /
Peter Lambrianakos (Reg. No. 58,279)
Lead Counsel for Patent Owner
FABRICANT LLP
411 Theodore Fremd Avenue,
Suite 206 South
Rye, New York 10580
Tel. 212-257-5797
Fax. 212-257-5796

CERTIFICATE OF SERVICE

A copy of the foregoing Patent Owner's Preliminary Response has been served on Petitioner's counsel of record as follows:

Lisa Nguyen
Email: lisa.nguyen@allenoverly.com
Email: AO_Meta_Jawbone@allenoverly.com
Eric Lancaster (*Pro Hac Vice* to be filed)
Email: eric.lancaster@allenoverly.com
ALLEN & OVERLY LLP
550 High Street
Palo Alto, California 94301

David Tennant
Email: david.tennant@allenoverly.com
Alan Billharz
Email: alan.billharz@allenoverly.com
Colby Davis
Email: Colby.davis@allenoverly.com
ALLEN & OVERLY LLP
1101 New York Avenue NW
Washington, D.C. 20005

Daniel Margolis (*Pro Hac Vice* to be filed)
Email: daniel.margolis@allenoverly.com
Grace Wang
Email: grace.wang@allenoverly.com
ALLEN & OVERLY LLP
1221 Sixth Avenue
New York, New York 10020

Attorneys for Meta Platforms, Inc.

December 11, 2023

By: /Peter Lambrianakos /
Peter Lambrianakos (Reg. No. 58,279)
Lead Counsel for Patent Owner

IPR2023-01291
PATENT NO. 8,467,543

FABRICANT LLP
411 Theodore Fremd Avenue,
Suite 206 South
Rye, New York 10580
Tel. 212-257-5797
Fax. 212-257-5796