

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ROKU, INC.,
Petitioner,

v.

UNIVERSAL ELECTRONICS, INC.,
Patent Owner.

IPR2022-00943
Patent 10,930,276 B2

Before PATRICK M. BOUCHER, SHARON FENICK, and
RUSSELL E. CASS, *Administrative Patent Judges*.

FENICK, *Administrative Patent Judge*.

DECISION
Granting Institution of *Inter Partes* Review
35 U.S.C. § 314, 37 C.F.R. § 42.4

I. INTRODUCTION

A. *Background and Summary*

Roku, Inc. (“Petitioner”) filed a petition for *inter partes* review challenging claims 1–16 (“challenged claims”) of U.S. Patent 10,930,276 B2 (Ex. 1001 (“the ’276 patent”)). Paper 2 (“Pet.” or “Petition”). Universal Electronics, Inc. (“Patent Owner”) timely filed a Preliminary Response. Paper 6 (“Prelim. Resp.”).

The Board has authority to determine whether to institute an *inter partes* review. *See* 35 U.S.C. § 314(b); 37 C.F.R. § 42.4(a). Under 35 U.S.C. § 314(a), we may not authorize an *inter partes* review unless the information in the Petition “shows that there is a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition.” For the reasons that follow, we institute an *inter partes* review as to the challenged claims of the ’267 patent on all grounds of unpatentability presented.

B. *Real Parties in Interest*

Petitioner identifies only itself as the real party in interest. Pet. 4.

Patent Owner identifies only itself as the real party in interest. Paper 4 (Patent Owner’s Mandatory Notices), 1.

C. *Related Matters*

Petitioner does not identify any related matters. Pet. 4. Patent Owner lists as related matters seven district court cases, one ITC case, and twenty IPR proceedings, and identifies several patent applications that are still pending and which are related to the ’276 patent by a claim of priority. Paper 4, 1–3.

D. The '276 Patent

The '276 patent is titled “Apparatus, System and Method for Directing Voice Input in a Controlling Device” and generally relates to a device controlling other appliances in an environment using speech input, for example to allow a user to use voice commands to control different types of consumer electronic devices. Ex. 1001, codes (54), (57), 1:29–33. The '276 patent describes, as an example, a situation in which a controlling device controls two or more smart appliances. *Id.* at 4:49–58. The controlling device may be, e.g., a standalone device, a tablet, PDA, cell phone, smart watch, computer, remote control, or intelligent digital assistant, among other possibilities. *Id.* at 4:50–63. The smart appliances may be, e.g., a TV, set top box, gaming system, or streaming device, among other media devices. *Id.* at 5:22–39.

The controlling device transmits a signal to a target device to communicate a command. *Id.* at 4:64–5:7. This may be in response to a voice command received by the controlling device from a user. *Id.* at 5:7–11. For example, the controlling device may have a soft key or a mechanical key that activates a coupled microphone, allowing the user’s voice to be used to determine the user’s intent to perform an action on an intended appliance. *Id.* at 5:45–58. Alternatively, the controlling device may continuously listen for a voice command. *Id.* at 5:66–6:10.

The '276 patent details a method that may be used by the controlling device. *Id.* at 13:5–14:50, Fig. 6. A user of a controlling device utters a voice command, which is received by the processor of the controlling device, and typically is stored by the controlling device. *Id.* at 13:16–24. The processor then determines an appliance the voice command relates to; this may be done at the controlling device or through consultation with a

local or remote device. *Id.* at 13:25–14:34. The controlling device formats the voice command into a data format associated with the identified appliance and transmits it to that device via a transmitter/transceiver. *Id.* at 14:35–50.

In one embodiment, the controlling device controls appliance-generated sound levels in an environment in connection with receiving speech input, as shown in Figure 7, reproduced below. *Id.* at 3:52–54.

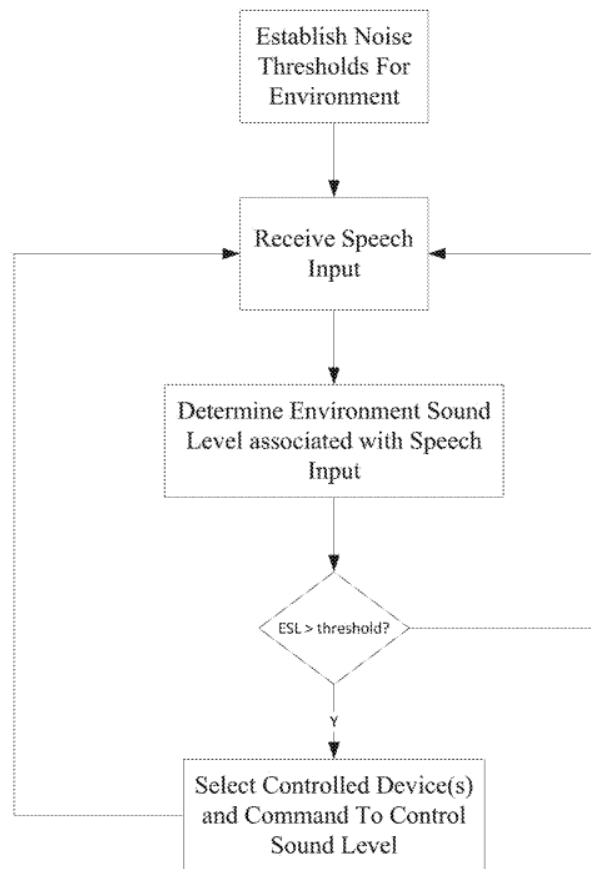


Figure 7

Figure 7 is a flow diagram that shows how one or more controlled devices may receive sound level commands in response to certain situations.

Id. at 3:52–54, 19:21–29, Fig. 7. In this embodiment, a predetermined threshold may be established by calibrating the controlling device. *Id.* at 19:36–41. A speech processing service associated with the controlling device performs a loudness analysis to determine a loudness of speech being provided to the controlling device and/or a loudness of an environment (e.g. background noise) in which the controlling device is operating. *Id.* at 19:21–29, 19:31–35. Based on the loudness estimation as compared to the predetermined threshold, the system can determine whether the speech input is being provided loudly, softly, or normally and whether the environment is loud, quiet, or normal. *Id.* at 19:41–57. Using this information, the controlling device may perform additional functions, like adjusting the volume of music or other output of a controlled device. *Id.* at 19:58–20:15. This may be done when an environment is loud and the user is attempting to provide speech communications to the controlling device. *Id.* at 20:16–36. If desired, the controlled device can be automatically returned to the prior state when the system determines that no further speech communications are likely to be issued by the user. *Id.* at 20:51–60.

E. Challenged Claims

Challenged claims 1 is independent, and all other challenged claims (2–16) depend directly or indirectly from claim 1.

Challenged claim 1 is reproduced below, with the addition of reference numbers in brackets, corresponding to the identification of limitations in the Petition.

1. A method for controlling a controllable appliance resident in an environment which includes a device adapted to receive speech input, comprising:

- [1.a] using a first sound data captured from the environment in which the device is operating to establish a noise threshold;

- [1.b] receiving at the device a speech input;
- [1.c] using a second sound data captured from the environment in which the device is operating at the time the speech input is received by the device to determine a noise level;
- [1.d] determining if the determined noise level is greater than the established noise threshold; and
- [1.e] in direct response to it being determined that the determined noise level is greater than the established noise threshold, causing one or more commands to be automatically issued to the controllable appliance to thereby cause the controllable appliance to transition from a first state having a first volume level to a second state having a second volume level that is less than the first volume level.

Ex. 1001, 22:36–56.

F. References Relied Upon

Petitioner relies upon the following prior art references:

Reference	Exhibit
Hart et al., US 9,251,787 B1, issued Feb. 2, 2016 (“Hart-787”)	1004
Fu et al., US 2017/0126192 A1, published May 4, 2017 (“Fu”)	1005
Rosenberg, US 9,509,269 B2, issued Nov. 29, 2016 (“Rosenberg”)	1006
Hart et al., US 9,466,286 B1, issued Oct.11, 2016 (“Hart-286”)	1007

Pet. 8.

Petitioner also relies on the Declaration of Mr. Bruce McNair.

(Ex. 1003).

G. Prior Art and Asserted Grounds

Petitioner asserts that the claims are unpatentable on the following grounds:

Claim(s) Challenged	35 U.S.C. § ¹	Reference(s)/Basis
1–11, 13–16	103	Hart-787, Fu, Rosenberg
12	103	Hart-787, Fu, Rosenberg, Hart-286

Pet. 8.

II. ANALYSIS

A. Legal Standards

“In an [*inter partes* review], the petitioner has the burden from the onset to show with particularity why the patent it challenges is unpatentable.” *Harmonic Inc. v. Avid Tech., Inc.*, 815 F.3d 1356, 1363 (Fed. Cir. 2016) (citing 35 U.S.C. § 312(a)(3) (requiring [*inter partes*] review petitions to identify “with particularity . . . the evidence that supports the grounds for the challenge to each claim”)). This burden never shifts to Patent Owner. *See Dynamic Drinkware, LLC v. Nat’l Graphics, Inc.*, 800 F.3d 1375, 1378 (Fed. Cir. 2015) (citing *Tech. Licensing Corp. v. Videotek, Inc.*, 545 F.3d 1316, 1326–27 (Fed. Cir. 2008)) (discussing the burden of proof in *inter partes* review).

Furthermore, Petitioner must explain with particularity how the prior art would have rendered the challenged claims unpatentable. 35 U.S.C. § 312(a)(3); 37 C.F.R. § 42.104 (“The petition must specify where each

¹ The Leahy-Smith America Invents Act (“AIA”), Pub. L. No. 112-29, 125 Stat. 284, 287–88 (2011), amended 35 U.S.C. § 103. The ’267 patent does not contain a claim of priority relating to a date before March 16, 2013 (the effective date of the relevant amendment), so the AIA version of § 103 applies.

element of the claim is found in the prior art patents or printed publications relied upon.”).

A claim is unpatentable under 35 U.S.C. § 103 if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious to a person having ordinary skill in the art to which said subject matter pertains. *See KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved on the basis of underlying factual determinations, including (1) the scope and content of the prior art; (2) any differences between the claimed subject matter and the prior art; (3) the level of skill in the art; and (4) objective evidence of nonobviousness, i.e., secondary considerations². *See Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966).

B. Level of Ordinary Skill in the Art

The level of skill in the art is a factual determination that provides a primary guarantee of objectivity in an obviousness analysis. *Al-Site Corp. v. VSI Int’l Inc.*, 174 F.3d 1308, 1323–1324 (Fed. Cir. 1999) (citing *Graham v. John Deere Co.*, 383 U.S. at 17–18; *Ryko Mfg. Co. v. Nu-Star, Inc.*, 950 F.2d 714, 718 (Fed. Cir. 1991)).

Petitioner argues that:

On or before 7/12/2017, [one of ordinary skill in the art] would have had a bachelor’s degree in computer science, in electrical engineering, or a related field, and approximately three years of experience or equivalent study in voice-controlled devices in smart home systems. Additional graduate education could

² In the present record, neither party presents arguments or evidence relating to secondary considerations.

substitute for professional experience, or significant experience in the field could substitute for formal education.

Pet. 18.

Patent Owner also uses July 12, 2017 as the priority date for its analysis, and argues that one of ordinary skill in the art as of that date:

would have had a bachelor's degree that involved software design and development coursework, for example, electrical engineering, computer engineering, computer science, industrial engineering, or a similar degree, and at least one year of work experience in software programming, development, or design of audio capture devices Additional education might substitute for some of the experience, and substantial experience might substitute for some of the educational background.

Prelim. Resp. 16 (citing Ex. 2001 (declaration of Dr. Michael S. Scordilis) ¶¶ 21–28). Dr. Scordilis asserts that Petitioner's definition "imports a number of specific and nuanced work-experience requirements that are not necessary to understand the invention of the '276 patent," but also notes that he performed his analysis from the point of view of a person of skill as defined by each party and "the differences between [the different levels of skill] did not affect my overall conclusions." Ex. 2001 ¶¶ 25, 28.

Based on the present record including the disclosure in the '267 patent, for the purposes of institution, we preliminarily adopt Petitioner's proposed definition of the level of ordinary skill in the art, which is consistent with the level of skill as reflected in the prior art. We find that this definition is consistent with the level of skill in the art as reflected by the prior art. *See Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001). Our decision here would not be different if we applied Patent Owner's proposed definition, in which we would substitute "approximately" for "at least" to avoid ambiguity.

C. Claim Construction

In interpreting the claims of the '276 patent, we “us[e] the same claim construction standard that would be used to construe the claim[s] in a civil action under 35 U.S.C. [§] 282(b).” 37 C.F.R. § 42.100(b) (2021). The claim construction standard includes construing claims in accordance with the ordinary and customary meaning of such claims as would have been understood by one of ordinary skill in the art and the prosecution history pertaining to the patent. *See id.*; *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–14 (Fed. Cir. 2005) (*en banc*).

“Importantly, the person of ordinary skill in the art is deemed to read the claim term not only in the context of the particular claim in which the disputed term appears, but in the context of the entire patent, including the specification.” *Phillips*, 415 F.3d at 1312. An inventor may rebut the presumption that the claim carries its ordinary and customary meaning by providing a definition of the term in the specification with reasonable clarity, deliberateness, and precision. *In re Paulsen*, 30 F.3d 1475, 1480 (Fed. Cir. 1994). In the absence of such a definition, limitations are not to be read from the specification into the claims. *In re Van Geuns*, 988 F.2d 1181, 1184 (Fed. Cir. 1993). Only those terms that are in controversy need be construed, and only to the extent necessary to resolve the controversy. *Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (citing *Vivid Techs., Inc. v. Am. Sci. & Eng'g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999)).

Petitioner, while arguing that no construction is necessary, asserts that the proper interpretation of “one or more commands” in claim 1 would include “conventional remote control commands.” Pet. 18–19. Patent

Owner does not request any specific claim construction or specifically dispute this assertion. Prelim. Resp. 16–17.

We need not construe this claim term at this point in the proceeding for the purpose of institution. *Nidec Motor Corp.*, 868 F.3d at 1017.

D. Asserted Obviousness Over Hart-787, Fu, and Rosenberg

Petitioner contends claims 1–11 and 13–16 would have been obvious over Hart-787, Fu, and Rosenberg. Pet. 26–58.

1. Hart-787 (Ex. 1004)

Hart-787 is titled “Altering Audio To Improve Automatic Speech Recognition” and relates to “altering audio being output by a voice-controlled device to enable more accurate automatic speech recognition (ASR) by the voice-controlled device.” Ex. 1004, codes (54), (57). Hart-787 describes that if speech is detected while a speech-recognizing device is outputting audio and the device determines that a user is going to provide a subsequent command, the device may attenuate or pause its audio output to facilitate speech recognition of the user’s subsequent command. *Id.* at code (57), 1:49–51. “By altering the output of the audio, an audio signal generated from the sound subsequently captured by the device will have less noise and, hence, will have a higher signal-to-noise ratio (SNR).” *Id.* at 2:7–10. According to Hart-787, the increased SNR will increase the accuracy of speech recognition and the chances that the device will be able to decode the user’s voice command. *Id.* at 2:10–13.

Hart-787 describes a system in which an electronic voice-controlled device is in a home environment with a user who may interact with the device. *Id.* at 3:54–58; Fig. 1. The voice-controlled device has a microphone and a speaker unit to facilitate audio interactions with the user. *Id.* at 4:15–18. The voice-controlled device has access to a speech

recognition device to recognize speech captured by the microphone, such as utterances by the user. *Id.* at 4:22–25, 5:2–6. The device includes a media player that may output content, including audio output via its speaker. *Id.* at 5:26–30, 8:34–37. Audio may also be output by a speaker of another device. *Id.* at 5:34–36.

An audio modification engine used by the device may be used to modify the output of the audio being output by the speaker of the device or the speaker of the other device in order to increase efficacy of speech recognition. *Id.* at 5:34–37. “For instance, in response to receiving an indication that the user . . . is going to provide a voice command to the device, the audio modification engine . . . may somehow modify the output of the audio to increase the accuracy of speech recognition performed on an audio signal generated from sound captured by the microphone.” *Id.* at 5:37–43. Hart-787 describes that this be done by attenuating (turning down), pausing, or any other alteration to the audio being output. *Id.* at 2:26–27, 5:47–51.

In an exemplary process flow described, the device captures sound within the environment, generates an audio signal based on the sound, and determines, from that audio signal, that the user is going to provide a command to the device. *Id.* at 7:66–8:1, 8:34–53, Fig. 3 (steps 302, 304, 306, and 308). The device then determines how to alter the audio being output under the control of the device and alters the audio accordingly. *Id.* at 8:54–59, Fig. 3 (steps 310, 312). This may include a determination that audio should be attenuated (turned down) based on a volume of a user’s utterance or a volume of the audio being output by the device. *Id.* at 9:18–25. For instance, it may alter the audio “based on a ratio between the user’s

utterance intended to gain the attention of the device (e.g. ‘wake up’) and the volume of the audio.” *Id.*

2. *Fu (Ex. 1005)*

Fu is titled “Method, Device, and Computer-Readable Medium for Adjusting Volume.” Ex. 1005, code (54). Fu relates to media player devices and to adjusting the volume of these devices automatically. *Id.* code (57), ¶¶ 2, 3, 5, 54. For example, a media player device may be playing media including a sound component. *Id.* ¶¶ 25–26, 53, 70, 77. The device captures ambient noise in the vicinity, such as via a microphone or decibel meter in the device or connected to the device. *Id.* ¶¶ 26–28, 53, 77. A volume adjustment may be made to the sound being output. *Id.* ¶¶ 29, 54. For example, the volume may be increased if the ambient environment is noisy or decreased if the ambient noise decreases. *Id.*

In certain embodiments, the volume of sound may be adjusted based on a predetermined threshold. *Id.* ¶¶ 33–37, Fig. 2. Fu describes that “[w]hen the amplitude of the ambient noise is large (i.e. exceeds the first predetermined threshold), the ambient environment is noisy, and thus the current volume of the sound may be increased automatically so as to improve the effect of playing sound.” *Id.* ¶ 37. Fu additionally describes that, “[w]hen the amplitude of the ambient noise changes from high to low, the current volume of the sound may be decreased automatically so as to improve the effect of playing sound. Such an implementation may easily improve the efficiency of adjusting volume.” *Id.*

3. *Rosenberg (Ex. 1006)*

Rosenberg is titled “Ambient Sound Responsive Media Player” and relates to capturing an ambient audio signal, determining if certain “characteristic forms” are present within that ambient audio signal, and, if

so, reducing the output of a media player from a first volume to a second volume. Ex. 1006 codes (54), (57), 2:16–44. Rosenberg explains that “[a] characteristic form is a sound or signal that when detected by the media player will cause an audible adjustment to the output of the media player.” *Id.* at 8:17–20. Some provided examples for characteristic forms are: recognition of a particular user’s voice; recognition of a particular verbal utterance such as a word, name, or phrase (e.g. pre-set utterances, or generic words commonly used to summon attention such as “help” or “excuse me”); and recognition of nonverbal sounds such as non-verbal human utterances, alarm/siren/emergency sounds, animal sounds, and characteristic environmental sounds. *Id.* at 7:20–61, 7:64–8:6, 8:21–27, 10:20–67.

One exemplary media player includes an audio subsystem with a microphone used for detection of sound signals from the user’s local ambient environment. *Id.* at 4:49–54, 5:62–6:18, Fig. 1 (elements 100, 95, 95A). The audio subsystem also includes headphones. *Id.* at 7:19–24, Fig. 1 (element 95B).

In one exemplary process flow, ambient audio signals are captured through the microphone and signal processing is performed on them to determine if a characteristic form is identified. *Id.* at 2:56–57, 6:66–7:22, 7:62–8:27, Fig. 2 (elements 201, 202, 203). If a characteristic form is identified, then the currently playing media audio signal is reduced in volume for a period of time so the user can better hear ambient sounds. *Id.* at 8:20–29, 8:32–54, 9:15–53, Fig. 2 (elements 203, 206, 208).

Rosenberg includes claim 5, dependent from Rosenberg’s claim 2, itself dependent from Rosenberg’s claim 1. Rewritten in independent form, this claim is directed to:

A method comprising:

capturing an ambient audio signal;

processing the ambient audio signal to determine whether one or more characteristic forms are present within the ambient audio signal;

automatically reducing an output of a media player from a first volume to a second volume gradually over a period of time if the one or more characteristic forms are present within the ambient audio signal, *wherein reducing the output further comprises reducing the output when one or more characteristic forms exceeds a volume threshold, wherein the volume threshold is determined based on the volume of the ambient audio signal*; and

maintaining the output of the media player at the second volume for a predetermined time before returning the output from the media player to the first volume, wherein the length of the predetermined time is based on the one or more characteristic forms, and wherein a first characteristic form is associated with a first length of predetermined time and a second characteristic form is associated with a second length of predetermined time different from the first length of predetermined time.

Id. at 13:66–14:21, 14:28–30 (language from claim 2 in italics, language from claim 5 in italics and underlined).

4. *Independent Claim 1*

Petitioner contends claim 1 would have been obvious over Hart-787, Fu, and Rosenberg. Pet. 19–45. Patent Owner presents arguments relating to claim 1’s obviousness over the proposed combination and relating to the propriety of the combination. Prelim. Resp. 17–34.

(a) *Preamble*

Petitioner contends that “to the extent this preamble is limiting” it is taught by Hart-787, which describes a method for altering the audio being

output from a controllable appliance resident in the same environment as a voice-controlled device. Pet. 31–34 (citing Ex. 1004, 3:66–4:3, 4:15–18, 5:43–51, Figs. 1, 2; Ex. 1003 ¶¶ 34–42, 94–96). Patent Owner does not present any argument with respect to this contention.

Without determining whether the preamble is limiting, on the present record, and for the purposes of institution, Petitioner has sufficiently shown that Hart-787, in the proposed combination of Hart-787, Fu, and Rosenberg, teaches or suggests the preamble of claim 1.

(b) *Limitation 1.a*

With respect to limitation 1.a, “using a first sound data captured from the environment in which the device is operating to establish a noise threshold,” Petitioner argues that a combination of Hart-787, Fu, and Rosenberg teaches this limitation. Pet. 34–41.

Describing its proposed combination, Petitioner argues that Hart-787 discloses a voice-controlled device with a microphone that detects ambient audio for the environment, including user commands, and that an audio modification engine in the voice-controlled device can reduce audio output of another device based on the ambient audio. *Id.* at 34–35 (citing Ex. 1004, 1:30–39, 2:7–13, 4:15–17, 5:34–43, 5:47–51, 8:39–42, Fig. 1; Ex. 1003 ¶ 98). Petitioner contends that Hart-787 teaches certain actions, including the attenuation of audio, being performed to increase the accuracy of speech recognition. *Id.* at 35–36 (citing Ex. 1004, 5:52, 8:54–63; Ex. 1003 ¶ 99). Petitioner argues that while Hart-787 does not specify how to determine when to attenuate the audio, one of ordinary skill would have applied Fu’s teachings regarding comparing current levels of ambient noise to a predetermined threshold to evaluate the volume of ambient noise. *Id.* at 36–

39 (citing *inter alia* Ex. 1005 ¶¶ 26, 37, 40–42, 45–47, 53, 54, Figs. 2–4, 6; Ex. 1003 ¶¶ 101–105).

Petitioner argues that the use of a predetermined dynamic threshold (based on ambient noise) is one of two possible options for use as a threshold in the proposed combination of Hart-787 and Fu, and it would have been obvious to implement such dynamic thresholds in light of Hart-787’s teachings. Pet. 30–31, 39–40 (citing Ex. 1003 ¶¶ 92, 106). Petitioner additionally cites Rosenberg’s teaching of a volume threshold determined based on the volume of the ambient audio signal in which it appears, taught in Rosenberg’s system where reducing an output of a media player occurs when a “characteristic form” (such as a detected trigger sound) exceeds a volume threshold determined based on the volume of ambient audio. Pet. 40–41 (citing Ex. 1006, 6:1–13, 7:62–8:6, 8:32–37, 10:62–67, claim 5 (13:66–14:21, 14:28–30); Ex. 1003 ¶¶ 107–108).

Petitioner contends that one of ordinary skill would have been motivated to modify the Hart-787 device to supply details regarding when to attenuate audio. *Id.* at 26–31, 40–41. Petitioner contends that one of ordinary skill would have included an adjustment to the audio output of a wirelessly-linked device when the voice-controlled device determined that the ambient noise level at the time of a wake word is greater than a threshold as in Fu, and, for the determination of that threshold given variable ambient noise, would have been motivated to further modify that combination to use a dynamic threshold that advantageously adjusts to the environment as in Rosenberg. *Id.* (citing Ex. 1005 code (57), ¶¶ 3, 24, 32, 54; Ex. 1006, 10:42–67; Ex. 1003 ¶¶ 87–93, 108–109). Petitioner further argues that one of ordinary skill would have had a reasonable expectation of success in doing this. *Id.* (citing Ex. 1003 ¶ 93).

Patent Owner presents several arguments relating to limitation 1.a. Prelim. Resp. 19–21. First, Patent Owner argues that Fu does not teach the *establishment* of a threshold, and that “F[u]’s *use* of a threshold is irrelevant because it is not dependent on ‘sound data captured from the environment in which the device is *operating*.’” *Id.* at 19. But, as discussed above, Petitioner relies on Rosenberg for details regarding establishment of a threshold related to the sound in the environment in the device is operating, and Fu does teach a predetermined threshold. *See* Pet. 39–41 (citing, *inter alia*, Ex. 1005 ¶ 37; Ex. 1006, claim 5 (13:66–14:21, 14:28–30); Ex. 1003 ¶¶ 107–108).

Patent Owner additionally argues that “[t]he threshold in the system of F[u] is entirely unrelated to the current audio output of the device, as the threshold of F[u] is based only on background noise levels.” Prelim. Resp. 19 (quoting (with added emphasis) limitation 1.a; citing Ex. 2001 ¶¶ 105–106). However, there is no requirement in claim 1 that the threshold be based on the current audio output of a device or that the current audio output of a device be included in the recited noise threshold or the recited sound data captured from the environment. This argument appears to be based on Dr. Scordilis’s testimony that “the claimed configuration” requires “*total* noise in an environment including ambient noises and noises generated by audio output of the system itself” to be analyzed; on the present record it is not clear what this assertion is based on. *See* Ex. 2001 ¶ 106.

Patent Owner also argues Rosenberg does not establish a threshold using sound data captured from the environment in which the device is operating, but rather uses measured loudness of a volume of a wake word, which Patent Owner contends is not analogous to the threshold limitation of claim 1. *Id.* at 20 (citing Ex. 2001 ¶ 103). Patent Owner contends that

“[t]he disclosure of Rosenberg . . . addresses the volume of a wake word [that] simply is not analogous to the ‘threshold’ limitation of claim 1.” *Id.* However, preliminarily, we find Petitioner’s contention that Rosenberg discusses establishing a volume threshold based on the volume of a captured ambient audio signal sufficiently shows the establishment of a noise threshold using sound data captured from the environment, as recited in limitation 1.a, in Petitioner’s proposed combination with the teachings of Hart-787 and Fu.

Patent Owner argues that Petitioner is mistaken in its contention that “in utilizing a noise threshold, there are only two options: pre-set and dynamically captured,” and that Petitioner has not provided any support for this statement. *Id.* at 21 (quoting Pet. 30). On the present record and for the purposes of institution, we agree that Petitioner has not sufficiently shown it will be likely to succeed in showing this, but we base our institution on Petitioner’s argument regarding teachings of Rosenberg and the assertions regarding known benefits of dynamic thresholds rather than on any implicit *KSR*-type argument by Petitioner that only two possible options existed. *See* Pet. 30–31 (citing *KSR v. Teleflex*, 550 U.S. at 398, 420 (2007) (“When there is a design need or market pressure to solve a problem and there are a finite number of identified, predictable solutions, a person of ordinary skill has good reason to pursue the known options within his or her technical grasp.”)).

Patent Owner additionally argues that Petitioner has not adequately shown a motivation for the combination of Hart-787, Fu, and Rosenberg. Prelim. Resp. 27–34. Patent Owner contends that Fu and Rosenberg are not in the same field of endeavor as Hart-787. *Id.* at 28. With respect to the proposed combination of elements from the references, Patent Owner argues

that Fu and Rosenberg both “focus on background noise that is distinct from audio output of the system.” *Id.* at 28–29 (citing Ex. 1005 ¶ 24; Ex. 1006, 3:12–18; Ex. 2001 ¶ 126). In contrast, Patent Owner argues that Hart-787 “suggests that the audio output of a device should be adjusted downward while a user is providing voice commands.” *Id.* at 28–29 (citing Ex. 1004, 2:22–37). Patent Owner notes that Fu raises audio output when environmental noise increases, and Rosenberg’s system is designed for use with noise-isolating headphones. *Id.* at 29–30 (citing Ex. 1006 4:36–40, 10:42–47; Ex. 2001 ¶¶ 127–128). Petitioner argues that these differences amount to Hart-787 “explicitly discredit[ing] the utility” of Rosenberg’s underlying technology. Patent Owner additionally argues that Petitioner’s proposed motivations rely on improper hindsight bias. *Id.* at 31–33.

On the present record and for the purpose of institution, we determine that Petitioner has adequately showed a proper combination of the prior art references. With respect to arguments regarding field of endeavor, we note that “[t]wo separate tests define the scope of analogous prior art: (1) whether the art is from the same field of endeavor, regardless of the problem addressed and, (2) if the reference is not within the field of the inventor’s endeavor, whether the reference still is reasonably pertinent to the particular problem with which the inventor is involved.” *In re Bigio*, 381 F.3d 1320, 1325 (Fed. Cir. 2004), quoted in *Donner Tech., LLC v. Pro Stage Gear, LLC*, 979 F.3d 1353, 1359 (Fed. Cir. 2020) (applying the tests from *In re Bigio* in an *inter partes* context). Each of the prior art references is an “audio monitoring method[]” (Prelim. Resp. 28) in which information received by audio-receiving microphones is used to attenuate what is played by a device. Ex. 1004, code (57); Ex. 1005, code (57); Ex. 1006, code (57).

On the present record, Petitioner has sufficiently shown that these references are in the same field of endeavor. *See* Pet. 61 (citing Ex. 1003 ¶ 159).

Additionally, based on the current record, we find Petitioner’s motivation to combine argument sufficient for institution. In particular, the distinction Patent Owner seeks to draw between the references does not undermine Petitioner’s argument and evidence that one of ordinary skill in the art would have, in order to effectuate Hart-787’s teaching of attenuating the audio output of a linked device if a user has indicated a command is to be issued, looked to Fu’s teaching of volume adjustment based on a predetermined threshold and Rosenberg’s teaching of a noise level threshold based on a measurement of environmental noise. Pet. 26–30. We additionally agree that Petitioner has, preliminarily and on this record, shown sufficiently that one of ordinary skill in the art would have been motivated to make the combination and that one of ordinary skill in the art would have had a reasonable expectation of success in doing so. *Id.* at 28–31 (Ex. 1003 ¶¶ 89, 91, 93). Notably, we find particularly compelling Hart-787’s express disclosure regarding the goal of improving SNR for spoken commands. Ex. 1004, 1:64–2:37 (cited at Pet. 28), 9:21–25 (describing attenuating the audio more if the ratio between the user’s utterance and the volume of the audio) (cited at Pet. 22, 27).

Certain of Patent Owner’s arguments regarding alleged differences in the focus of the prior art references and incompatibilities in their purposes appear akin to arguing bodily incorporation, rather than arguments regarding what the combined teachings of Hart-787, Fu, and Rosenberg would have suggested to one of ordinary skill in the art. *See, e.g., In re Keller*, 642 F.2d 413, 425 (CCPA 1981) (“The test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the

primary reference. . . . Rather, the test is what the combined teachings of those references would have suggested to those of ordinary skill in the art.”). And the cited differences between Hart-787 and Rosenberg, on the present record, do not show that Hart-787 discredits Rosenberg or teaches away from it. *See In re Fulton*, 391 F.3d 1195, 1201 (Fed. Cir. 2004) (prior art does not teach away from claimed subject matter merely by disclosing a different solution to a similar problem unless the prior art also criticizes, discredits, or otherwise discourages the solution claimed).

On this record, Petitioner sufficiently shows that the combination of Hart-787, Fu, and Rosenberg teaches limitation 1.a of claim 1.

(c) *Limitation 1.b*

Petitioner contends that limitation 1.b is taught by Hart-787, which teaches a microphone detecting audio at the device. Pet. 41–42 (citing Ex. 1004, 4:15–18, Fig. 1; Ex. 1003 ¶¶ 110–111). Patent Owner does not present any argument with respect to this contention.

On the present record, and for the purposes of institution, Petitioner has sufficiently shown that Hart-787, in the proposed combination of Hart-787, Fu, and Rosenberg, teaches limitation 1.b of claim 1.

(d) *Claim 1 – Limitation 1.c*

Petitioner contends that Hart-787, in combination with Fu and Rosenberg, teaches using a second sound data captured from the environment in which the device is operating at the time the speech input is received to determine a noise level, as recited in limitation 1.c of claim 1. Pet. 42. Petitioner notes that when the device in Hart-787 receives a speech input (such as “wake up”), it receives that input along with other environmental ambient noise, and uses that received audio signal to

determine a noise level. *Id.* (citing Ex. 1004, 5:58–59, 8:39–42, 9:18–21; Ex. 1003 ¶¶ 57, 112–114).

Patent Owner argues that, in Hart-787, “once a voice command is recognized, it is irrelevant what the ambient noise level is — Hart-787 simply indicates that the output volume is always reduced.” Prelim. Resp. 22. Patent Owner additionally argues that Hart-787 does not teach or suggest discretely capturing sound data multiple times, and therefore does not teach or suggest “second sound data.” *Id.* at 22–23.

On the current record and for the purposes of institution, Petitioner has sufficiently shown that Hart-787 in the proposed combination teaches or suggests using a second sound data captured from the environment in which the device is operating at the time speech input is received by the device. Hart-787 teaches that “a microphone . . . capture[s] the sound and generate[s] the audio signal, with the sound includ[ing] both the audio being output and additional audio (e.g., speech from a user, background noise, etc.).” Ex. 1004, 8:39–42 (cited at Pet. 42). While Patent Owner argues that Hart-787 always reduces the output volume and that “it is irrelevant what the ambient noise level is” (Prelim. Resp. 22), this disregards Hart-787’s discussion that the attenuation effected may be based on the volume of the audio being played by the controlled device in Hart-787. Ex. 1004, 9:18–25 (cited at Pet. 22, 27, 29, 31, 36, 39–40, 42); *see id.* at 2:66–33 (describing that the device may “attenuate the audio [being output by the device] a great deal” in certain situations and “a small amount (or not at all)” in other situations). And with respect to Patent Owner’s argument that Hart-787 does not teach capturing sound data multiple times, the combination presented uses Fu and Rosenberg to teach or suggest a predetermined threshold based on ambient sound, and does not rely only on teachings of

Hart-787. *See In re Keller*, 642 F.2d at 425–426 (finding, in an examination context, that “one cannot show non-obviousness by attacking references individually where . . . the rejections are based on combinations of references.”). We determine that on the present record, for the purposes of institution, there is a reasonable likelihood that Petitioner will prevail in showing the limitations requiring a first and second sound data capture in the combination it proposes.

For these reasons, on the present record, and for the purposes of institution, Petitioner has sufficiently shown that Hart-787, in the proposed combination of Hart-787, Fu, and Rosenberg, teaches limitation 1.c of claim 1.

(e) *Claim 1 – Limitations 1.d and 1.e*

Petitioner contends that the combination of Hart-787, Fu, and Rosenberg, teaches limitations 1.d and 1.e of claim 1. Pet. 43–45. Petitioner argues limitation 1.d, “determining if the determined noise level is greater than the established noise threshold,” with reference back to its arguments with respect to claim 1.a, relating to Rosenberg’s threshold being used in Hart-787’s device modified to apply Fu’s method of determining whether to adjust the volume of audio output by comparing ambient noise volume to a predetermined threshold. Pet. 43 (citing Ex. 1003 ¶¶ 115–116). Limitation 1.e recites that, “in direct response” to a determination (in 1.d) that the determined noise level is greater than the established noise threshold, “one or more commands” are caused “to be automatically issued to the controllable appliance” to cause it to transition from a first state with a first volume level to a second state with a second volume level less than the first volume level. Ex. 1001, 22:49–56. Petitioner argues that this is taught or suggested by Hart-787’s audio modification engine, which may attenuate the

audio if a speech input is being received. Pet. 43–44 (citing Ex. 1004, code (57), 1:64–2:7, 2:24–27, 5:34–51, 8:56–59; Ex. 1003 ¶¶ 44–45, 118).

Petitioner additionally argues that Fu’s teachings in the proposed combination teach that the audio modification would be in direct response to a determination that an environmental noise level is greater than an established noise threshold. *Id.* at 44 (citing Ex. 1003 ¶ 119). While Petitioner acknowledges that certain examples in Fu disclose increasing volume output (rather than decreasing it) in response to a determination that the volume of ambient noise exceeds a predetermined threshold, Petitioner contends that Fu does not require such an increase, and cites Rosenberg’s teachings regarding volume reduction in playing volume of media content when a determination has been made that the ambient noise includes certain triggers that indicate possibly relevant information in ambient noise for the user. *Id.* at 44–45 (citing Ex. 1006, 9:15–10:20, 12:42–67; Ex. 1003 ¶ 120).

Patent Owner argues that Rosenberg’s environmental noise threshold “is limited to a threshold used to ensure that identified sound signatures captured from the ambient environment are sufficiently loud as to be used as trigger sounds,” and does not involve a threshold generated based on previously-captured sound data. Prelim. Resp. 23–34 (citing Ex. 2001 ¶¶ 117–118). Patent Owner additionally argues that Fu discloses a pre-defined threshold for determining when a background noise level is loud enough to warrant a change in audio volume. *Id.* at 24 (citing Ex. 2001 ¶¶ 117–118). Thus, Patent Owner argues, neither Rosenberg nor Fu teaches using first sound data to establish a noise level (as recited in limitation 1.a) and second sound data to make the determination recited in limitation 1.d. *Id.* at 23–24. Patent Owner argues that Hart-787 only teaches adjusting output volume every time a voice command is received (Prelim. Resp. 26).

These arguments are addressed to the references individually and not to what the combination would have suggested to one of ordinary skill in the art, as addressed previously with reference to limitation 1.c., *supra* at II.D.4.(d).

Patent Owner contends that Fu teaches only increasing volume when ambient noise surpasses a pre-established threshold, while limitation 1.e of claim 1 recites the opposite. Prelim. Resp. 25–27 (citing Ex. 1005 ¶¶ 24, 77; Ex. 2001 ¶¶ 121–122). However, contrary to Patent Owner’s assertion (Prelim. Resp. 26–27), Petitioner relies on Rosenberg for its teaching that speakers may be decreased in response to a determination of a characteristic form has been identified in the ambient noise, as well as to Fu’s more general examples of adjusting by increasing or decreasing the volume. *See* Pet. 44–45 (citing Ex. 1005 ¶ 29; Ex. 1006, 9:15–10:20; Ex. 1003 ¶ 120).

For the reasons discussed above, on the current record and for the purposes of institution, Petitioner has sufficiently shown that the combination of Hart-787, Fu, and Rosenberg teaches or suggests limitations 1.d and 1.e of claim 1.

(f) *Conclusion – Claim 1*

For the reasons discussed above, based on the record before us, we determine Petitioner’s offered arguments, evidence, and supporting testimony establish a reasonable likelihood that Petitioner would prevail in showing that claim 1 would have been obvious over the combination of Hart-787, Fu, and Rosenberg.

5. *Claims 2–11 and 13–16*

As noted above, Petitioner also challenges dependent claims 2–11 and 13–19 using the combination of Hart-787, Fu, and Rosenberg and provides a detailed analysis of how the prior art disclosures teach the elements of those challenged claims. Pet. 45–58. Other than as discussed above, Patent

Owner does not additionally challenge Petitioner's analysis in its Preliminary Response.

We have reviewed Petitioner's arguments and the underlying evidence cited in support and are persuaded that, at this stage, Petitioner sufficiently demonstrates a reasonable likelihood of succeeding in its challenges to claims 2–11 and 13–16.

E. Asserted Obviousness Over Hart-787, Fu, Rosenberg, and Hart-286

Petitioner challenges claim 12 as obvious over a combination of Hart-787, Fu, Rosenberg, and Hart-286 and provides a detailed analysis of how the prior art disclosures teach the elements of those challenged claims and the motivation to combine Hart-286 with Hart-787's device as modified by Fu and Rosenberg. Pet. 58–63. Other than as discussed above, Patent Owner does not additionally challenge Petitioner's analysis in its Preliminary Response.

We have reviewed Petitioner's arguments and the underlying evidence cited in support and are persuaded that, at this stage, Petitioner sufficiently demonstrates a reasonable likelihood of succeeding in its challenge to claim 12.

III. ORDER

For the reasons given, it is:

ORDERED that pursuant to 35 U.S.C. § 314(a), an *inter partes* is *instituted* on claims 1–16 of the '276 patent with respect to all grounds set forth in the Petition; and

FURTHER ORDERED that pursuant to 35 U.S.C. § 314(a), an *inter partes* review of the '276 patent is instituted commencing on the entry date of this Order, and pursuant to 35 U.S.C. § 314(c) and 37 C.F.R. § 42.4, notice is given of the institution of a trial.

IPR2022-00943
Patent 10,930,276 B2

FOR PETITIONER:

Scott A. McKeown
Matthew J. Rizzolo
Brendan F. McLaughlin
ROPES & GRAY LLP
scott.mckeown@ropesgray.com
matthew.rizzolo@ropesgray.com
brendan.mclaughlin@ropesgray.com

FOR PATENT OWNER:

S. Benjamin Pleune
Thomas W. Davison
Ryan W. Koppelman
Dan O'Connor
Nicholas T. Tsui
ALSTON & BIRD LLP
ben.pleune@alston.com
tom.davison@alston.com
ryan.koppelman@alston.com
dan.oconnor@alston.com
nick.tsui@alston.com