



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/946,701	11/15/2010	Jason S. Flaks	330080-US-NP	1467
69316	7590	12/26/2019	EXAMINER	
MICROSOFT CORPORATION ONE MICROSOFT WAY REDMOND, WA 98052			TARKO, ASMAMAW G	
			ART UNIT	PAPER NUMBER
			2482	
			NOTIFICATION DATE	DELIVERY MODE
			12/26/2019	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

chriochs@microsoft.com
usdoCKET@microsoft.com

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Ex parte JASON S. FLAKS and AVI BAR-ZEEV

Appeal 2018-002460
Application 12/946,701
Technology Center 2400

Before MAHSHID D. SAADAT, CARL L. SILVERMAN, and
ALEX S. YAP, *Administrative Patent Judges*.

YAP, *Administrative Patent Judge*.

DECISION ON APPEAL

Appellant appeals under 35 U.S.C. § 134(a) from the Examiner's Final rejection of claims 1–20.¹ (Final Act. 1 (Final Office Action, mailed May 5, 2017, “Final Act.”).) We have jurisdiction under 35 U.S.C. § 6(b), and we heard the appeal on October 24, 2019.

We AFFIRM IN PART.

¹ We use the word “Appellant” to refer to “applicant” as defined in 37 C.F.R. § 1.42. Appellant identifies Microsoft Technology Licensing, LLC. as the real party in interest. (Appeal Br. 3.)

STATEMENT OF THE CASE

Introduction

According to the Specification, the claimed invention relates to “system and method providing semi-private conversation using an area microphone between one local user in a group of local users and a remote user.” (November 15, 2010 Specification (“Spec.”) Abstract.) Claim 1 is illustrative, and is reproduced below (with minor reformatting):

1. A method of providing a semi-private conversation between a local user and a remote user, comprising:
 - receiving by way of each of an array of plural microphones, both of a first voice output from a respective first local user and a second voice output from a respective second local user, said first and second local users being situated in a sound-sharing first physical environment in which the first and second local users are in a two-way sound cross-talk relationship with one another and in which the microphone array is also situated;
 - using the microphone array to localize a respective origin position in the first physical environment for each of the first and second received voice outputs, the respective origin positions corresponding to respective locations of the first and second local users in the first physical environment;
 - associating the localized voice outputs with the correspondingly located first and second local users;
 - using the microphone array to isolate the respective voice output of a selected one of the first and second local users from other sounds present in the sound-sharing first physical environment; and
 - directing the isolated voice output of the selected local user to a selected third user in a second

physical environment, where sounds in the first physical environment are not propagated to the second physical environment.

Prior Art and Rejections on Appeal

The following table lists the prior art relied upon by the Examiner in rejecting the claims on appeal:

Name	Reference	Date
Wang	US 2002/0086733 A1	July 4, 2002
Danieli et al. (“Danieli”)	US 2003/0216178 A1	November 20, 2003
Mao et al. (“Mao”)	US 2006/0239471 A1	October 26, 2006
Hunter	US 2007/0211067 A1	September 13, 2007
Tashev et al. (“Tashev”)	US 2008/0288219 A1	November 20, 2008
Thakkar et al. (“Thakkar”)	US 2009/0210491 A1	August 20, 2009

Claims 1–3, 5, 7, 8, 10–15, and 17–19 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Thakkar in view of Mao. (*See* Final Act. 3–7.)

Claim 4 stands rejected under 35 U.S.C. § 103(a) as being unpatentable over Thakkar in view of Mao and Hunter. (*See* Final Act. 7.)

Claim 6 stands rejected under 35 U.S.C. § 103(a) as being unpatentable over Thakkar in view of Mao and Wang. (*See* Final Act. 8.)

Claims 9 and 16 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Thakkar in view of Mao and Tashev. (*See* Final Act. 8–9.)

Claim 20 stands rejected under 35 U.S.C. § 103(a) as being unpatentable over Thakkar in view of Danieli. (*See* Final Act. 9–10.)

ANALYSIS

We have reviewed the Examiner’s rejection in light of Appellant’s arguments that the Examiner has erred. We are not persuaded that the Examiner erred in rejecting claims 1–6, 9, and 15 but are persuaded that the Examiner erred in rejecting claims 7, 8, 10–14, and 16–20.

Claim Term “semi-private”

The term “semi-private” appears in the preamble² of independent claims 1 (semi-private conversation) and 15 (semi-private communications). Appellant contends that the term “semi-private” should be construed to “exclude[] totally public conversations where the utterances of any participant in a conference is always heard by all other participant[s] in the conference.” (Appeal Br. 9.) We agree with the Examiner that Appellant’s construction is not the broadest reasonable construction of the term. (Ans. 11–12.) Specifically, we agree with the Examiner that the Specification:

mainly discusses two people on sound sharing environment communicating with two other people in the remote sound sharing environment and the discussion is mainly between one person in the local environment with the another person of the remote environment without need for wearing special headsets or utterance isolating private microphones. As described by the appellant’s disclosure[,] the voice of a prospective user will be isolated using a microphone array and directed to the remote user using a speaker array and vice versa. It is not clear to person of

² We do not opine on whether this term in the preamble is limiting.

having ordinary skilled in the art how the speaker array placed in the open sound sharing environment can direct the sound of the local user specifically to one of the two remote user. Appellant have failed to particularly claim and clearly show the means of performing what is been argued.

(*Id.* at 11.) While paragraphs 3 and 18 of the Specification (cited by Appellant for support of its construction) state that “[t]he semi-private conversation experience is provided without the use of traditional sound isolating technology, such as microphones and head-sets,” and that **“[d]irectional transmission technology** may be used to output the local user’s utterances to the remote user in the remote environment,” they do not support Appellant’s proposed construction that the term “exclude[] totally public conversations where the utterances of any participant in a conference is always heard by all other participant in the conference.” (Appeal Br. 9.) In other words, the directional output transmission of a local user’s utterances to a remote user does not mean that the utterance is not heard by other remote users in that conference room. The Specification merely states that such utterance would be “directed to” a remote user “without the use of traditional sound isolating technology, such as microphones and head-sets.” (Spec. ¶¶ 3, 18.) For the foregoing reasons, we decline Appellant’s invitation to import a negative limitation into the term “semi-private.”

Claim Term “using the microphone array to localize a respective origin position in the first physical environment . . .”

For the “using the microphone array to localize a respective origin position in the first physical environment . . .” limitation of claim 1, the Examiner finds that while Thakkar teaches “meeting participant can share

conference room or participate from remote location,”³ it “fail[s] to show explicitly that using a microphone array to localize and isolate the respective voice output of a selected one of the first and second local users from other sounds present in the sound-sharing first physical environment.” (Final Act. 4–5.) However, the Examiner notes that “Mao shows using a microphone array to localize and isolate the respective voice output of a selected one of the first and second local users from other sounds present in the sound-sharing first physical environment” and that “[i]t would have been obvious to person of having ordinary skilled in the art to combine the voice localization and isolation method of Mao into the system of Thakkar to improve the smooth communication between players who are in sound-sharing environment in a local and remote location.” (*Id.*)

Appellant contends that:

Fig[ure] 2 [of] *Thakkar* is not teaching that the voice of participant 154-2 is being isolated by use of the microphones array from the voices of participant 154-1 and 154-p of the same room 150, but rather that the image of participant 154-2 (his ‘image chunk’) is being ‘annotated’ by adding that participant's name 270-1-d next to his image.

(Appeal Br. 12–15.)

First, as discussed above, the Examiner is using the localization and isolation method of Mao. (Final Act. 5.) Second, we agree with the Examiner that “Thakkar identifies the users based on their prospective environment as 154-1, 154-2 . . . 154-n where 154 shows the environment the user is located and the number after the hyphen identify the user” and “using letters or numbers to identify a user is within the ordinary skill of

³ The Examiner cites to paragraphs 41 and 46 and Figure 2 of Thakkar. (Final Act. 4.)

person of having ordinary skilled in the art.” (Ans. 12.) We also agree with the Examiner that “[i]t would have been obvious to person of having ordinary skilled in the art to combine the voice localization and isolation method of Mao into the system of *Thakkar* to improve the smooth communication **between** players who are in sound-sharing environment in a local and remote location.” (Final Act. 5; *Thakkar*, ¶¶ 43 (“the media analysis module 210 may detect a number of participants 154-1-*p* using both image analysis and voice analysis on audio content from the input media streams 204-1-*f*: Other types of media content may be used as well.”), 59 (“the media analysis module **210** may output location information for each region in the media frames **252-1-g** that are used to form the image chunks with detected participants. . . .”); Mao, Fig. 1, ¶¶ 12, 15, 88.)

Appellant also contends that “location module 232 of *Thakkar* tracks the location of a user within a video frame so that annotations may be applied to the frame” but “the location module 232 takes as input media frames *that are associated with audio content but do not themselves include audio content*; the location module 232 therefore does not use audio content as an input.” (Reply Br. 2, emphasis added.) *Thakkar*, however, discloses that “the media analysis module **210** may detect a number of participants **154-1-*p*** using both image analysis *and voice analysis on audio content from the input media streams 204-1-*f**.” (*Thakkar*, ¶ 43, emphasis added.) Importantly, even *assuming arguendo* that the input media frames that are associated with audio content but do not themselves include audio content, the combination would teach the limitation at issue and one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. *See In re Merck & Co.*, 800 F.2d

1091, 231 USPQ 375 (Fed. Cir. 1986); *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981).

We are similarly not persuaded by Appellant’s contention that “Thakkar does not disclose using the techniques performed by the location module 232 for locating origin positions of voice outputs rather than images of speakers.” (Reply Br. 3; Appeal Br. 12–15; *see also* Appeal Br. 20–21 (discussing same issue for claim 2).) An image of a location is representative of that location and locating participants using voice analysis on audio content from the input media streams is akin to locating participants in that location. Furthermore, “Mao shows using a microphone array to localize and isolate the respective voice output of a selected one of the first and second local users from other sounds present in the sound-sharing first physical environment” and one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. (Final Act. 5.) *See In re Merck & Co.*, 800 F.2d at 231; *In re Keller*, 642 F.2d at 413.

Appellant further contends that Thakkar does not “provide technical details that would enable one of ordinary skill in the art to modify the location module 232 to locate origin positions of sounds.” (Reply Br. 3.) This argument is waived because this argument is raised for the first time on Reply. 37 C.F.R. § 41.41(b)(2). “[A]n issue not raised by an [A]ppellant in its opening brief . . . is waived.” *Cf. Optimus Tech., Inc. v. Ion Beam Appl’ns S.A.*, 469 F.3d 978, 989 (Fed. Cir. 2006) (citations and internal quotations omitted); *Cf. McBride v. Merrell Dow and Pharms., Inc.*, 800 F.2d 1208, 1211 (D.C. Cir. 1986) (internal citations omitted) (“Considering an argument for the first time in a reply brief . . . is not only unfair to the

appellee but also entails the risk of an improvident or ill-advised opinion on the legal issues tendered.”).

Claim Term “directing the isolated voice output of the selected local user to a selected third user in a second physical environment . . .”

For the “*directing the isolated voice output of the selected local user to a selected third user in a second physical environment . . .*” limitation of claim 1, the Examiner finds that while Thakkar teaches “meeting participant can share conference room or participate from remote location,” it “fail[s] to show explicitly that using a microphone array to localize and isolate the respective voice output of a selected one of the first and second local users from other sounds present in the sound-sharing first physical environment.” (Final Act. 4–5.) However, the Examiner notes that “Mao shows using a microphone array to localize and isolate the respective voice output of a selected one of the first and second local users from other sounds present in the sound-sharing first physical environment” and that “[i]t would have been obvious to person of having ordinary skilled in the art to combine the voice localization and isolation method of Mao into the system of Thakkar to improve the smooth communication between players who are in sound-sharing environment in a local and remote location.” (*Id.*)

Appellant contends that “that *Mao* does not teach or suggest *directing* an isolated voice output of a selected local user to a selected third (remote) user so as to provide ‘semi-private’ conversations.” (Appeal Br. 20.) According to Appellant, Mao does “not teach or suggest ‘semi-private’ conversations that include the selective *directing* of an isolated voice output to a selected third user, just the elimination of noise sources at the input

side.” (*Id.*) The Examiner points out that “Mao is presented to cure a deficiency of Thakkar[,] which is voice isolation of a prospective user when other voices are present in the room [and] Mao also shows the remote users on the screen labeling person 102’ and person 112.” (Ans. 13.) In other words, the Examiner is not relying on Mao alone for “directing” the isolated voice output of the selected local user to a remote third user. Appellant appears to suggest that “directing to,” similar to its proposed construction for “semi-private,” requires excluding the isolated utterances from being heard by all other participant in the remote location. This, however, is not the broadest reasonable reading of the claim. The claim only requires the isolated voice output of the selected local user be “directed to” a remote third user. It does not required exclusion of all other participants in the remote location from hearing the isolated voice output. Even using “directional transmission technology,” which according to the Specification “may be used” to direct the isolated voice output of the selected local user to a remote third user, (Spec. ¶ 3) anyone in the room within a reasonable distance from the third user would still be able to hear the voice output.

*“tracking user location by detecting user location in
a field of view of a depth image camera”*

Claim 3 recites, in part, “tracking user location by detecting user location in a field of view of a depth image camera.” The Examiner points to paragraphs 48 and 63 for this limitation. (Final Act. 5.) Appellant contends that “neither of these paragraphs teaches [n]or suggests a depth image camera” but does not explain why. (Appeal Br. 21.) *See In re Lovin*, 652 F.3d 1349, 1357 (Fed. Cir. 2011) (“[W]e hold that the Board reasonably interpreted Rule 41.37 to require more substantive arguments in an appeal

brief than a mere recitation of the claim elements and a naked assertion that the corresponding elements were not found in the prior art.”); *cf. In re Baxter Travenol Labs.*, 952 F.2d 388, 391 (Fed. Cir. 1991) (“It is not the function of this court to examine the claims in greater detail than argued by an appellant, looking for [patentable] distinctions over the prior art.”) Therefore, we sustain the Examiner’s rejection of this claim.

“determining a conversational relationship between the selected local user and another user in the second physical environment.”

Claim 5 recites, in part, “determining a conversational relationship between the selected local user and another user in the second physical environment.” The Examiner points to paragraph 27 for this limitation. (Final Act. 5.) Appellant, however, argues that while the paragraph “does mention the possibility of multiple conferences [and] that it says nothing about ‘semi-private’ conversations (Claim 1) or the locations of the participants involved in respective semi-private conversations.” (Appeal Br. 21.) As discussed above, however, these arguments are not persuasive.

Therefore, in view of the foregoing, we sustain the Examiner’s rejection of independent claim 1 and claims 2–6, 9, and 15, which are not argued separately.

Claims 7 and 8

Claim 7 recites, in part, “receiving isolated utterances from the third user in the second physical environment and routing the utterances to the selected local user in the first physical environment” and claim 8 recites, in part, “routing comprises providing isolated utterances from the third user to a directional output aimed at the selected local user.” For these claims, the Examiner finds that they “have limitations similar to those treated in the

above rejections, and are met by the references as discussed above and are rejected for the same reasons of obviousness as used above.” (Final Act. 6.) Appellant, however, contends that “such hand waving does not comply with the notice requirements of 35 USC §132 [leaving t]he Board and Appellant to speculate on what constitutes the ‘similarity’ between Claims 7 [and] 8” (Appeal Br. 21–22; Reply Br. 4–7.) The Examiner responds that “[c]laims 7 and 8 are not presented to have a significant limitations different from what has been discussed above in that it only adds a third user while the prior art says it can be used with multiple users.” (Ans. 13–14.)

We agree with Appellant that the Examiner has not complied with the notice requirement of 35 USC §132. While the Final Office Action discusses how the prior art teaches isolating utterance in a multiuser environment and directing of such isolated utterance to a participant in a remote location, the Examiner, on this record, does not make any further findings relating to the other limitations of these claims. In the event of further prosecution, the Examiner may consider whether, in view of the prior art, it would have been obvious to a person of ordinary skill in the art to route isolated utterances from a user in a remote location (*i.e.*, the third user) to a selected local user (claim 7) and whether it would have been obvious to use a directional output aimed at the selected local user (claim 8). Therefore, in view of the foregoing, we do not sustain the Examiner’s rejection of claims 7 and 8.

Claims 10–14

The Examiner notes that “[s]ystem claims 10–12 are drawn to the system corresponding to the method using same as claimed in claims 1–3. Therefore system claims 10–12 corresponding to method claims 1-3, and is

rejected for the same reasons of obviousness as used above.” (Final Act. 6.) Appellant contends that independent “[c]laim 10 recites the ‘directional audio output device’ while method Claim 1 does not [and n]either of *Thakkar* and *Mao* have been shown to teach or suggest such a directional audio output device.” (Appeal Br. 22.) The Examiner does not address Appellant’s contention and we agree with Appellant that independent claim 10 contains at least one limitation that is not present in independent claim 1 and the Examiner, on this record, does not make any further findings relating to this limitation. In the event of further prosecution, the Examiner may consider whether, in view of the prior art, it would have been obvious to a person of ordinary skill in the art to use a directional output to direct audio output from a remote user to a localized source location of the first user. Therefore, in view of the foregoing, we do not sustain the Examiner’s rejection of independent claim 10 and dependent claims 11–14, which depend on claim 10.

Claims 15–20

The Examiner notes that “[c]laim 15 has limitations similar to those treated in the rejections of claim 1 above, and are met by the references as discussed above. Therefore, claim 15 is rejected for the same reasons of obviousness as used above.” (*Id.*) Appellant argues that “that [c]laim 15 recites the separate routings of the isolated utterances of the first user and of the isolated utterances of the second user while Claim 1 does not. Thus at least some of the terms of Claim 15 are rendered superfluous.” (*Id.* at 23.) Appellant’s argument is not persuasive because we agree with the Examiner that “the prior art says it can be used with multiple users.” (Ans. 14.)

Therefore, it would have been obvious to one of ordinary skill in the art to route utterance between different users in different environments.

Claim 16 recites, in part, “wherein the steps of isolating includes combining spatial filtering with regularization on the input to provide an isolated output.” The Examiner finds that “[c]laim 16 has limitations similar to those treated in the rejection of claim 9 above, and are met by the references as discussed above. Therefore, claims 16 is rejected for the same reasons of obviousness as used above.” (Final Act. 9.) Appellant contends that the Examiner’s rejection “does not comply with the notice requirements of 35 USC § 132 and instead leaves the Board and Appellant to mere speculation.” (Appeal Br. 24.) We agree and do not sustain the rejection of claim 16.

Claim 9 recites, in part, “wherein the step of localization includes combining spatial filtering with regularization on the received voice output to thereby provide at least two corresponding outputs.” Therefore, unlike claim 16, which recites providing “and isolated output,” claim 9 requires providing “at least two corresponding outputs.” The Examiner’s rejection for claim 9 is as follows:

Claim 9 also recites, but Thakkar and Mao failed to show the step of localization includes combining spatial filtering with regularization on the received voice output to thereby provide at least two corresponding outputs. However, Tashev shows step of localization includes combining spatial filtering with regularization on the received voice output to thereby provide at least two corresponding outputs are known (0027). It would have been obvious to person of having ordinary skilled in the art to localize the voice output by combining spatial filtering with regularization on the received voice output to thereby provide at least two corresponding outputs as shown in Tashev into the

systems of Thakkar and Mao to increase the efficiency and accuracy of voice filtration and directing in the right direction.

(Final Act. 9.) The Examiner, however, does not explain how Tashev teaches providing “an isolated output” or whether one of both of the “at least two corresponding outputs” is “an isolated output.” Therefore, in view of the foregoing, we do not sustain the Examiner’s rejection of claim 16 and claims 17–20, which depend on claim 16.

CONCLUSION

In summary:

Claim(s) Rejected	35 U.S.C. §	Basis/Reference(s)	Affirmed	Reversed
1–3, 5, 7, 8, 10–15, and 17–19	103(b)	Thakkar and Mao	1–3, 5, and 15	7, 8, 10– 14, and 17–19
4	103(b)	Thakkar, Mao, and Hunter	4	
6	103(b)	Thakkar, Mao, and Wang	6	
9 and 16	103(b)	Thakkar, Mao, and Tashev	9	16
20	103(b)	Danieli		20
Overall Outcome			1–6, 9, and 15	7, 8, 10– 14, and 16–20

Appeal 2018-002460
Application 12/946,701

DECISION

We affirm the Examiner's decision to reject claims 1–6, 9, and 15.
We reverse the Examiner's decision to reject claims 7, 8, 10–14, and 16–20.

No time period for taking any subsequent action in connection with this appeal may be extended under 37 C.F.R. § 1.136(a)(1)(iv).

AFFIRMED IN PART