Paper No. 21

UNITED STATES PATENT AND TRADEMARK OFFICE

_____

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

_____

Ex parte JOONYOUL MAENG and ERROL R. WILLIAMS

_____

Appeal No. 2000-0546
Application 08/509,228

_____

ON BRIEF

_____

Before KRASS, FLEMING, and GROSS, **Administrative Patent Judges.**

FLEMING, **Administrative Patent Judge.**

DECISION ON APPEAL

This is a decision on appeal from the final rejection of claims 1-30, all the claims pending in the instant application.

The invention relates to an automatic voice tracking camera system, an automatic camera controller and method for automatically controlling a camera to track the position of a speaker using the speaker's voice. The system (12) includes a camera(18), a microphone array (14), a beamformer (30), and a camera controller (36). See Appellants' specification on page 6,

lines 13-26, page 9, lines 13-28 and associated figures 1-2. The automatic camera controller (16) includes a camera controller (36) and an interface (38). See Appellants' specification on page 8, lines 15-16 and 30-36 and associated figure 2.

The independent claims 1, 12, and 20 present in the application are reproduced as follows:

1. An automatic voice tracking camera system, comprising:

a first camera operable to receive control signals for controlling a view of the first camera;

a microphone array comprising a plurality of microphones, the microphone array operable to receive a voice of a speaker and to provide an audio signal representing the voice;

a beamformer coupled to the microphone array, the beamformer operable to receive the audio signal, to generate from the audio signal speaker position data representing a position of the speaker as coordinates for a point in space, and to provide the speaker position data; and

a camera controller coupled to the beamformer and to the first camera, the camera controller operable:

to receive the speaker position data,

to determine an appropriate responsive camera movement based upon the coordinates for the point in space,

to generate camera control signals based upon the appropriate responsive camera movement, and

to provide the camera control signals to the first camera such that the view of the first camera automatically tracks the position of the speaker.

12. An automatic camera controller for automatically

controlling a camera to track a position of a speaker, the automatic camera controller comprising:

a camera controller operable:

to receive speaker position data representing the position of the speaker as coordinates for a point in space,

to determine an appropriate responsive camera movement based upon the coordinates for the point in space,

to generate camera control signals based upon the appropriate responsive camera movement, and

to provide the camera control signals to at least one camera such that a view of the at least one camera automatically tracks the position of the speaker; and

an interface coupled to the camera controller, the interface operable to communicate with a host video conference system and with the camera controller.

20. A method of automatically controlling a camera to track a position of a speaker using a voice of the speaker, comprising:

receiving a voice of the speaker and providing an audio signal representing the voice;

processing the audio signal to generate speaker position data representing the position of the speaker as coordinates for a point in space;

determining an appropriate responsive camera movement from the speaker position data based upon the coordinates for the point in space;

generating camera control signals based upon the appropriate responsive camera movement; and

providing the camera control signals to a first camera such that a view of the first camera automatically tracks the position of the speaker.

References

The references relied on by the Examiner are as follows:

Ashida et al. (Ashida)	5,206,721	Apr. 27, 1993
Kannes	5,382,972	Jan. 17, 1995
Washino et al. (Washino)	5,625,410	Apr. 29, 1997 (filed Apr. 7, 1995)
Baker	5,686,957	Nov. 11, 1997 (filed Jan. 30, 1995)

Rejections at Issue

Claims 1, 4, 12, 14, 20 and 30 stand rejected under 35 U.S.C. § 102 as being anticipated by Baker. Claims 2-3, 5-11, 13, 15-19 and 21-27 stand rejected under 35 U.S.C. § 103 as being unpatentable over Baker and Ashida.[1] Claim 28 stands rejected under 35 U.S.C. § 103 as being unpatentable over Baker and Washino. Claim 29 stands rejected under 35 U.S.C. § 103 as being unpatentable over Baker and Kannes.

Rather than repeat the arguments of Appellants or the Examiner, we make reference to the Briefs[2] and the Answer for the respective details thereof.

---

[1] Appellants and Examiner mistakenly include claim 14 in the § 103 rejection of Baker in view of Ashida. In fact, claim 14 has only been rejected under § 102 by Baker.

[2] Appellants filed an appeal brief on March 4, 1999, Paper No. 17. On July 29, 1999, Appellants filed a reply brief, Paper No. 19, in response to the Examiner's answer, Paper No. 18, mailed June 8, 1999. The Examiner entered the reply brief and mailed notification, Paper No. 20, to Appellants on August 12, 1999.

OPINION

With full consideration being given the subject matter on appeal, the Examiner's rejections and the arguments of Appellants and Examiner, for the reasons stated infra, we reverse the Examiner's rejection of claims 1, 4, 12, 14, 20 and 30 under 35 U.S.C. § 102, and we reverse the Examiner's rejection of claims 2-3, 5-11, 13, 15-19 and 21-29 under 35 U.S.C. § 103.

We first will address the rejection of claims 1, 4, 12, 14, 20 and 30 under 35 U.S.C. § 102. "It is axiomatic that anticipation of a claim under § 102 can be found only if the prior art reference discloses every element of the claim[.]" ***See In re King***, 801 F.2d 1324, 1326, 231 USPQ 136, 138 (Fed. Cir. 1986) and ***Lindemann Maschinenfabrik GMBH v. American Hoist & Derrick Co.,*** 730 F.2d 1452, 1458, 221 USPQ 481, 485 (Fed. Cir. 1984).

Appellants argue that Baker does not disclose the speaker position data representing the position of the speaker as coordinates for a point in space. Appeal Brief, page 5, lines 9-16. More specifically, Appellants argue that the limitation, ""a point in space[,]" must be defined in three dimensions." Reply Brief, page 3, lines 10-11 and Appeal Brief, page 5, lines 22-23. Appellants assert that the generated data disclosed by

Baker is distinguishable from the three dimensional speaker position data generated by Appellants' invention. Appeal Brief, page 5, lines 20-24.

As pointed out by our reviewing court, we must first determine the scope of the claim. "[T]he name of the game is the claim." ***In re Hiniker Co.,*** 150 F.3d 1362, 1369, 47 USPQ2d 1523, 1529 (Fed. Cir. 1998).

We note that Appellants' claim 1 recites "the beamformer operable . . . to generate from the audio signal speaker position data representing a position of the speaker as coordinates for a point in space" and Appellants' claim 20 recites "processing the audio signal to generate speaker position data representing the position of the speaker as coordinates for a point in space." Appeal Brief, page 11, lines 8-11 and page 16, lines 6-8. Appellants' claim language, "speaker position data representing a position of the speaker as coordinates for a point in space" or "speaker position data representing the position of the speaker as coordinates for a point in space," reasonably allows for the reading of claims 1 and 20 to require the generation of speaker position data from an audio signal representing the position of the speaker as coordinates of a point in space.

We also note that Appellants' claim 12 has a different scope than claims 1 and 20. Claim 12 recites "a camera controller operable . . . to *receive* speaker position data representing the position of the speaker as coordinates for a point in space . . . and to provide the camera control signals to at least one camera such that a view of the at least one camera automatically tracks the position of the speaker" (emphasis added). Appeal Brief, page 14, lines 4-6 and 10-13. Appellants' claim language, "to receive speaker position data representing the position of the speaker as coordinates for a point in space," reasonably allows for the reading of claim 12 to require a camera controller to receive data that represents the position of the speaker as coordinates for a point in space and to automatically track the position of the speaker based upon the data received.

When interpreting a claim, words of the claim are generally given their ordinary and accustomed meaning, unless it appears from the specification or the file history that they were used differently by the inventor. ***Carroll Touch, Inc. v. Electro Mechanical Sys., Inc.*** 15 F.3d 1573, 1577, 27 USPQ2d 1836, 1840 (Fed. Cir. 1993). Although an inventor is indeed free to define the specific terms used to describe his or her invention, this must be done with reasonable clarity, deliberateness, and

precision. ***In re Paulsen,*** 30 F.3d 1475, 1480, 31 USPQ2d 1671, 1674 (Fed. Cir. 1994). Our reviewing court states in ***In re Zletz,*** 893 F.2d 319, 321, 13 USPQ2d 1320, 1322 (Fed. Cir. 1989) that "claims must be interpreted as broadly as their terms reasonably allow."

We note that Appellants' specification states that "speaker position data 34 comprises Cartesian coordinates defining the location of the speaker." Specification, page 9, lines 19-21. Additionally, we note that Appellants' specification describes the Cartesian coordinates in terms of three variables (x, y and z). Specification, page 19, line 36. Finally in referencing page 9 and Figure 9 of Appellants' specification, Appellants admit that their "invention uses beamforming to establish a speaker's location in three dimensions[.]" Appeal Brief, page 5, lines 22-23. Thus, for the reasons stated above, Appellants have shown that the speaker position data representing a position of the speaker as coordinates for a point in space must include coordinates in three dimensions.

The Examiner argues that Baker discloses speaker position data representing a position of the speaker as coordinates for a point in space. To support this position, the Examiner refers to

a passage of Baker which discusses determining the speakers' position by sampling audio signals in order to detect the largest amplitude signal. Column 9, lines 29-33. This signal information is then sent to an audio direction processor to determine the direction to steer the camera.

Upon careful review, we fail to find that Baker discloses speaker position data representing the position of the speaker as coordinates in three dimensions. Baker discloses gathering positional information about the speaker based on the strength of the amplitude of the auido signals. Based on this information, Baker discloses determining the direction to steer the camera by correlating between the location of the microphone and the video imaging system. Column 9, lines 29-38. As such, the informational data disclosed by Baker does not represent the position of the speaker in the form of coordinates having three dimensions, but rather represents a correlation between the microphone and video imaging system.

The Examiner also argues that Baker discloses speaker position data representing the position of the speaker as coordinates for a point in space since "[i]t is also possible, by normal audio beam steering techniques, to select points between microphones[.]" Column 10, lines 19-21. The Examiner states

that audio signals propagate from a point in space to a microphone and "without knowing the coordinates for a point in space of the speaker," steering cannot work. Examiner's Answer, page 14, line 8-14.

Upon consideration, we fail to find that this section of Baker discloses speaker position data representing a position of the speaker as coordinates for a point in space. This portion of Baker describes an alternative arrangement that steers the camera to a point between the microphones, rather than at the microphone. The steering location, whether at the microphone or at a point between the microphones, in Baker remains predefined by a correlation process and does not relate to the position of the speaker.

Lastly, the Examiner argues that Baker discloses the use of three-dimensional cameras to obtain three dimensional images, establishing the desirability for the speaker's location in three-dimensions. Examiner's Answer, page 15, lines 4-7. We find this argument misguided for several reasons. First, there is no discussion in Baker to generate speaker position as coordinates in three dimensions in order to assist in the three dimensional imaging taught by Baker. Second, speculating that the three dimensional camera requires that the beamformer in

Baker generate speaker position data as three dimensional data does not establish an element of a claim. Third, this argument is directed to modifying Baker to obtain speaker position data. This is a determination under obviousness and not a determination under anticipation as that which is before us.

Thus, we fail to find that Baker discloses "a beamformer . . . operable . . . to generate from the audio signal speaker position data representing a position of the speaker as coordinates for a point in space" as recited in claim 1 or the step of processing an audio signal "to generate speaker position data representing the position of the speaker as coordinates for a point in space" as recited in claim 20. Appellants have shown that speaker position data must be three dimensional and represents the position of the speaker as a point in space.

Since claim 12 does not recite the generation of data to represent the position of the speaker as coordinates for a point in space, we will treat the claim separately. As discussed above Baker does not disclose generating speaker position data having coordinates in three dimensions. Additionally, Baker provides no disclosure of a camera controller operable to receive speaker position data representing the position of the speaker as coordinates for a point in space or operable to provide camera

control signals to a camera such that the camera automatically tracks the position of the speaker based upon the coordinates for the point in space. Thus, we also fail to find that Baker discloses a camera controller operable "to receive speaker position data representing the position of the speaker as coordinates for a point in space . . . to provide the camera control signals to at least one camera such that a view of the least one camera automatically tracks the position of the at speaker" as recited in claim 12.

Therefore, we find that Baker fails to teach all of the limitations of claims 1, 12, 14, 20 and 30, and thus the claims are not anticipated by Baker.

We now turn to the rejection of claims 2-3, 5-11, 13, 15-19, and 21-29 under 35 U.S.C. § 103. The Examiner has not relied on the secondary references, Ashida, Washino or Kannes, to teach or suggest the elements in claims 1, 12 or 20 missing from Baker. As such, we also cannot sustain the rejections made under 35 U.S.C. § 103.

REMAND

After a careful review of the record, we remand the decision to the Examiner for further consideration.

We note that Appellants make specific reference to U.S. Patent Application 09/399,427, now U.S. Patent 5,737,431 to Brandstein et al., in addressing the operation of a microphone array and beamformer. Appellants' Specification, Page 9, lines 5-9. Brandstein discloses an automatic voice tracking camera system used in the teleconferencing environment that includes a beamformer connected to a camera in order to automatically steer the camera to the position of the speaker based on the speaker's voice. See column 1, line 66 through column 2, line 4 and 18-32, column 3, lines 27-39, column 6, lines 9-14, column 7, line 64 through column 8, line 1 and associated figures 1 and 3. In addition, Brandstein describes the beamformer connected to a microphone array to generate speaker position data defined in Cartesian coordinates (x, y and z directions). See column 9, line 61 - column 10, line 7, column 6, lines 15-24, and column 8, lines 16-23 and 44-54. However, Brandstein does not disclose the specifics of a camera controller operable to perform the limitations recited in claims 1, 12 and 20.

Baker also discloses an automated camera tracking system used in the teleconferencing environment that includes a camera, a microphone array, a beamformer, a camera controller and an interface. See column 9, lines 5-29, column 10, lines 10-16,

column 15, lines 44-50 and associated figures 1A and 5. Baker teaches a camera controller that receives data from the beamformer and sends signals based on this data to the camera in order to steer the camera in the direction of the speaker. See column 9, lines 20-23 and 29-33.

The Examiner should consider a rejection under 35 U.S.C. § 103 combining Baker and Brandstein. Factual findings regarding the beamformer recited in claim 1 operable to generate speaker position data in three dimensions or the step of generating the speaker position data recited in claim 20, as taught by Brandstein, should be addressed. Also, suggestions from Brandstein to use the speaker position data with a camera controller, such as the controller of Baker, should be considered. In addition, the Examiner should include factual findings addressing a camera controller operable to receive speaker position data as recited in claim 12 and as suggested by

Brandstein and Baker. Finally, the difference in scope as to the remaining claims should also be addressed, and appropriate rejections should be formulated.

REVERSED and REMANDED

ERROL A. KRASS Administrative Patent Judge	)	
MICHAEL R. FLEMING Administrative Patent Judge	)	BOARD OF PATENT APPEALS AND INTERFERENCES
ANITA PELLMAN GROSS Administrative Patent Judge	)	

MRF:pgg

Baker and Botts
2001 Ross Avenue
Dallas, TX 75201-2980