

S25, the distance from sensing device 20 (i.e., camera 22) to the person may be specified based on distance information (i.e., actually measured distance value) obtained from distance sensor 23.

[0050]

5 As a result of steps S24 and S25 described above, as shown in FIG. 7, human detection unit 32 can detect the location of a person in space 70. FIG. 7 schematically shows a result of detecting the location of a person (i.e., a view of table 40 from above). In the example of FIG. 7, the locations of three people A, B, and C are
10 detected. Note that human detection unit 32 detects the three-dimensional locations of the people (i.e., the three-dimensional coordinates of the people) but only needs to detect at least the two-dimensional locations of the people as shown in FIG. 7 (the directions represented by angles in the top view). In the following embodiment,
15 the locations of the people detected by human detection unit 32 will also be referred to as "first locations".

[0051]

 Human detection unit 32 can track the locations (i.e., the first locations) of people by repeating the operation in FIG. 6 at each unit
20 time. At this time, ID information may be assigned only once in step S23.

[0052]

[Operation of Utterance Amount Estimation Unit]

 Now, an operation of utterance amount estimation unit 34 will
25 be described more in detail. FIG. 8 is a flowchart of the operation of utterance amount estimation unit 34.

[0053]

 Utterance amount estimation unit 34 obtains the first locations detected by human detection unit 32 and the second location detected
30 by sound source detection unit 31 (S31). The first and second locations obtained at this time are detected at substantially the same time. The "substantially the same time" may include some time lags.

[0054]

 Next, utterance amount estimation unit 34 converts the first
35 locations represented by the three-dimensional coordinates into two-