

[DESCRIPTION]

[Title of Invention] ATTENTION DETERMINATION SYSTEM,
COMMUNICATION ANALYSIS SYSTEM, AND ATTENTION
DETERMINATION METHOD

5 [Technical Field]

[0001]

The present invention relates to an attention determination
system, a communication analysis system, and an attention
determination method.

10 [Background Art]

[0002]

It is important in an organization, such as a company, that
employees work on their tasks, while closely communicating with each
other. Patent Literature (PTL) 1 discloses, as a technique related to
15 such communication, an information providing device that provides
the psychological state of a subject as information to evaluate
intellectual activities at a conference or a meeting.

[Citation List]

[Patent Literature]

20 [0003]

[PTL 1] Japanese Unexamined Patent Application Publication No.
2004-112518

[Summary of Invention]

[Technical Problem]

25 [0004]

The present invention provides an attention determination
system, for example, capable of determining how much a subject pays
attention to communication.

[Solutions to Problem]

30 [0005]

An attention determination system according to an aspect of
the present invention includes: a human detection unit that obtains
video information about a video including people within a space in
which a display and a whiteboard are placed, and detects first
35 locations which are locations of the people in the space, based on the

video information obtained; a head orientation detection unit that obtains the video information, and detects an orientation of a head of a subject among the people, based on the video information obtained; and an attention determination unit that determines whether the subject is paying attention to communication among the people in the space, based on the first locations detected and the orientation of the head of the subject as detected. When one of the people other than the subject is speaking, the attention determination unit determines, as a period in which the subject pays attention to the communication, a period in which the head of the subject detected faces any of (i) one of the people other than the subject, (ii) the display, and (iii) the whiteboard.

[0006]

A communication analysis system according to an aspect of the present invention includes: the attention determination system described above; and a communication analysis unit that analyzes a quality of communication among the people in the space, based on whether the subject is paying attention determined.

[0007]

An attention determination method according to an aspect of the present invention includes: obtaining video information about a video including people within a space in which a display and a whiteboard are placed, and detecting first locations which are locations of the people in the space, based on the video information obtained; obtaining the video information, and detecting an orientation of a head of a subject among the people, based on the video information obtained; and determining whether the subject is paying attention to communication among the people in the space, based on the first locations detected and the orientation of the head of the subject detected. The determining whether the subject is paying attention includes determining a period in which the head of the subject detected faces any of one of the people other than the subject, the display, and the whiteboard, as a period in which the

subject pays attention to the communication, when one of the people other than the subject is speaking.

[0008]

5 A program according to an aspect of the present invention is for causing a computer to execute the attention determination method described above.

[Advantageous Effects of Invention]

[0009]

10 An attention determination system, for example, according to the present invention can determine how much a subject pays attention to communication.

[Brief Description of Drawings]

[0010]

[FIG. 1]

15 FIG. 1 is a block diagram showing a functional configuration of a communication analysis system according to an embodiment.

[FIG. 2]

FIG. 2 shows an example space to which the communication analysis system according to the embodiment is applied.

20 [FIG. 3]

FIG. 3 is an external view of a sensing device according to the embodiment.

[FIG. 4]

25 FIG. 4 is a flowchart of an operation of a sound source detection unit included in the communication analysis system according to the embodiment.

[FIG. 5]

FIG. 5 schematically shows a result of detecting the location of a sound source.

30 [FIG. 6]

FIG. 6 is a flowchart of an operation of a human detection unit included in the communication analysis system according to the embodiment.

[FIG. 7]

FIG. 7 schematically shows a result of detecting the location of a person.

[FIG. 8]

FIG. 8 is a flowchart of an operation of an utterance amount estimation unit included in the communication analysis system according to the embodiment.

[FIG. 9]

FIG. 9 schematically shows a result of detecting the location of a speaker.

10 [FIG. 10]

FIG. 10 shows example information indicating the amounts of utterance.

[FIG. 11]

FIG. 11 is a flowchart of an operation of a head orientation detection unit included in the communication analysis system according to the embodiment.

[FIG. 12]

FIG. 12 shows a three-dimensional coordinate space for illustrating an operation of the head orientation detection unit included in the communication analysis system according to the embodiment.

[FIG. 13]

FIG. 13 illustrates correction of an angle of a head.

[FIG. 14]

FIG. 14 is a flowchart of an operation of an attention determination unit included in the communication analysis system according to the embodiment.

[FIG. 15]

FIG. 15 illustrates a target direction of a subject.

30 [FIG. 16]

FIG. 16 shows example information indicating periods of attention.

[FIG. 17]

FIG. 17 is a flowchart of an operation of a communication analysis unit included in the communication analysis system according to the embodiment.

[FIG. 18]

5 FIG. 18 shows an example display screen showing, on an information terminal, a score representing the quality of communication.

[Description of Embodiment]

[0011]

10 Now, an embodiment will be described in detail with reference to the drawings. The embodiment described below is a mere general or specific example. The numerical values, shapes, materials, elements, the arrangement and connection of the elements, steps, step orders etc. shown in the following embodiment are thus mere
15 examples, and are not intended to limit the scope of the present invention. Among the elements in the following embodiment, those not recited in the independent claims will be described as optional.

[0012]

20 The figures are schematic representations and not necessarily drawn strictly to scale. The same reference signs represent substantially the same configurations in the drawings and redundant description may be omitted or simplified.

[0013]

[Embodiment]

25 [Configuration]

 First, a configuration of a communication analysis system according to the embodiment will be described. FIG. 1 is a block diagram showing a functional configuration of the communication analysis system according to the embodiment. FIG. 2 shows an
30 example space to which the communication analysis system is applied.
[0014]

 As shown in FIGS. 1 and 2, communication analysis system 10 is used in an office with space 70, such as a conference room, to analyze the quality of communication among people within space 70.
35 Space 70 is a closed space, for example, but may be an open space.

Examples of space 70 include, in addition to a conference room, an open rest area within an office space (e.g., a part of an office space with chairs and tables). Space 70 is not necessarily physically divided, but may be divided in a whole space by illumination light or an air flow. For example, an office space lit in day light with a color temperature of 5000 K includes, at a corner, a warm color area with a color temperature of 3000 K, which may be regarded as space 70. [0015]

Communication analysis system 10 includes sensing device 20 and information processing system 30. First, sensing device 20 will be described with reference to FIG. 3 in addition to FIGS. 1 and 2. FIG. 3 is an external view of sensing device 20. In FIG. 3, (a) is a top view of sensing device 20 and (b) is a side view of sensing device 20. Note that FIG. 3 shows no distance sensor 23. [0016]

Sensing device 20 is placed on table 40 in space 70 to sense sound and videos in space 70. Specifically, sensing device 20 is placed at the center of table 40. Sensing device 20 includes microphone array 21 and a plurality of cameras 22 and distance sensor 23. [0017]

Microphone array 21 obtains sound in space 70 and outputs sound information (i.e., a plurality of sound signals) about the obtained sound. Specifically, microphone array 21 includes a plurality of microphone elements each of which obtains sound in space 70 and outputs sound signals related to the obtained sound. [0018]

Each of cameras 22 captures a video (i.e., a movie) of people within space 70 and outputs video information about the video. Each camera 22 is a general camera with a CMOS image sensor but may also be a fisheye camera. Including four cameras to capture the whole surroundings of sensing device 20 from the top of table 40, sensing device 20 only needs to include at least one camera capable of capturing all the people within space 70. [0019]

Distance sensor 23 measures the distance from sensing device 20 (e.g., one of cameras 22) to a subject and outputs distance information indicating the measured distance to the subject. The subject is a person within space 70, for example. Distance sensor 23 is a Time of Flight (TOF) Light Detection and Ranging (LiDAR) device, for example, but may be a distance image sensor or any other suitable device. While sensing device 20 only needs to include at least one distance sensor 23, sensing device 20 may include a plurality of distance sensors 23 corresponding to cameras 22.

10 [0020]

Now, information processing system 30 will be described. Information processing system 30 establishes wired or wireless communication with sensing device 20. Information processing system 30 analyzes the quality of communication, based on sensing information (specifically, sound information, video information, and distance information) obtained from sensing device 20 through the communication. Information processing system 30 is an edge computer in a facility with space 70, for example, but may be a cloud computer outside the facility. If information processing system 30 is an edge computer, sensing device 20 and information processing system 30 may serve as an integral device. Some of the functions of information processing system 30 may be implemented by an edge computer and the others may be implemented by a cloud computer.

[0021]

Specifically, information processing system 30 includes sound source detection unit 31, human detection unit 32, head orientation detection unit 33, utterance amount estimation unit 34, attention determination unit 35, communication analysis unit 36, and storage unit 37.

30 [0022]

Sound source detection unit 31 obtains sound information about the sound obtained in space 70 from sensing device 20, and detects a second location, which is the location of the sound source in space 70, based on the obtained sound information.

35 [0023]

Human detection unit 32 obtains video information about a video showing people within space 70 from sensing device 20, and detects first locations, which are the locations of the people within space 70, based on the obtained video information.

5 [0024]

Head orientation detection unit 33 obtains video information about a video showing people within space 70 from sensing device 20, and detects the orientation of the head (i.e., the orientation of the face) of a person, based on the obtained video information. Head orientation detection unit 33 may detect the direction of the eyes of a person, based on the obtained video information.

[0025]

Utterance amount estimation unit 34 estimates the amount of utterance of a person, based on the first locations detected by human detection unit 32 and the second location detected by sound source detection unit 31.

[0026]

Attention determination unit 35 determines whether a person is paying attention to communication among people including the person within space 70, based on the following information. The information includes the first locations detected by human detection unit 32 and the orientation of the head of a person detected by head orientation detection unit 33.

[0027]

Communication analysis unit 36 analyzes the quality of communication, based on at least one of the following information. The information includes the amount of utterance of a person estimated by utterance amount estimation unit 34 and whether a person is paying attention determined by attention determination unit 35. Communication analysis unit 36 also outputs a result of analysis.

[0028]

Each of sound source detection unit 31, human detection unit 32, head orientation detection unit 33, utterance amount estimation unit 34, attention determination unit 35, and communication analysis unit 36 described above is a microcomputer or a processor. The

functions of sound source detection unit 31, human detection unit 32, head orientation detection unit 33, utterance amount estimation unit 34, attention determination unit 35, and communication analysis unit 36 are fulfilled, for example, by the microcomputer or the processor
5 executing computer programs stored in storage unit 37.

[0029]

Storage unit 37 is a storage device storing information necessary for executing the computer programs and fulfilling the functions of the elements. Storage unit 37 is a hard disk drive (HDD),
10 for example, but may be a semiconductor memory.

[0030]

Note that the system including sound source detection unit 31, human detection unit 32, and utterance amount estimation unit 34 will also be referred to as "speaker diarisation system 38". That is,
15 speaker diarisation system 38 includes sound source detection unit 31, human detection unit 32, and utterance amount estimation unit 34. Speaker diarisation system 38 may further include head orientation detection unit 33 or attention determination unit 35.

[0031]

20 The system including human detection unit 32, head orientation detection unit 33, and attention determination unit 35 will also be referred to as "attention determination system 39". That is, attention determination system 39 includes human detection unit 32, head orientation detection unit 33, and attention determination unit
25 35. Attention determination system 39 may further include sound source detection unit 31 or utterance amount estimation unit 34.

[0032]

[Operation of Sound source detection unit]

Now, an operation of sound source detection unit 31 will be
30 described more in detail. FIG. 4 is a flowchart of the operation of sound source detection unit 31.

[0033]

First, sound source detection unit 31 obtains sound information about the sound obtained in space 70 from microphone array 21 of
35 sensing device 20 (S11). Specifically, sound source detection unit

31 obtains a plurality of sound signals output by the plurality of microphone elements included in microphone array 21.

[0034]

Here, the obtained sound signals are time domain signals.

5 Sound source detection unit 31 Fourier-transforms the sound signals from the time domain signals to frequency domain signals (S12).

[0035]

Next, sound source detection unit 31 obtains a spatial correlation matrix from an input vector determined by the sound signals transformed into the frequency domain signals (S13). Here, the input vector $x(\omega, t)$ is expressed by the following equation, where microphone array 21 includes M microphone elements among which the m-th sound signal after the Fourier-transform has a vector $X_m(\omega, t)$. T means transpose.

15 [0036]

[Math. 1]

$$x(\omega, t) = [X_1(\omega, t) \cdots, X_M(\omega, t)]^T$$

[0037]

Spatial correlation matrix R is expressed by the equation below.

20 H means conjugate transpose. In the following description, the index ω of the frequency will be omitted for the sake of simplicity.

[0038]

[Math. 2]

$$R = E[x(t)x^H(t)]$$

25 [0039]

Sound source detection unit 31 then performs eigenvalue decomposition on the spatial correlation matrix described above to obtain eigenvectors (S14). Specifically, sound source detection unit 31 performs the eigenvalue decomposition on the spatial correlation matrix described above based on the following equation to obtain eigenvalue vectors $e_1, \dots, e_M$ and eigenvalues $\lambda_1, \dots, \lambda_M$.

[0040]

[Math. 3]

$$Re_n = \lambda_n e_n$$

[0041]

After that, sound source detection unit 31 detects the location
5 of the sound source from the eigenvectors (S15). Specifically, sound
source detection unit 31 can specify the volumes and directions of
sounds based on the eigenvectors, and detect the direction of a
relatively large sound as the direction (i.e., the location) of the sound
source.

10 [0042]

As a result, as shown in FIG. 5, sound source detection unit 31
can detect in which direction (i.e., at which angle) the sound source
is located with reference to location O of sensing device 20. FIG. 5
schematically shows a result of detecting the location of the sound
15 source (i.e., a view of table 40 from above). In the example of FIG.
5, two sound sources S1 and S2 are detected. Note that sound
source detection unit 31 only needs to detect at least the two-
dimensional location of a sound source as shown in FIG. 5 (the
directions represented by an angle in the top view), but may detect
20 the three-dimensional location of a sound source. In the following
embodiment, the location of a sound source detected by sound source
detection unit 31 will also be referred to as a "second location".

[0043]

Sound source detection unit 31 can track the location (i.e., the
25 second location) of the sound source by repeating the operation in
FIG. 4 at each unit time. Note that the "sound source" is specifically
a speaker (i.e., a person) within space 70 but may be a device within
space 70.

[0044]

30 [Operation of Human Detection Unit]

Now, an operation of human detection unit 32 will be described
more in detail. FIG. 6 is a flowchart of the operation of human
detection unit 32. While the following operation is actually

performed for video information obtained from each of four cameras 22, the following description assumes that video information is obtained from one camera 22, for the sake of simplicity.

[0045]

5 First, human detection unit 32 obtains video information about a video obtained in space 70 from camera 22 of sensing device 20 (S21).

[0046]

10 Next, human detection unit 32 specifies an area including any person within the video, based on the obtained video information (S22). Human detection unit 32 can specify the area including any person within the video by a technique of pattern matching or a machine learning model.

[0047]

15 Human detection unit 32 then assigns ID information to the specified area (i.e., the area including any person)(S23). For example, human detection unit 32 specifies three areas and assigns ID information A, B, and C to the three specified areas. In the following description, a person within area A will be referred to as
20 "person A", a person within area B will be referred to as "person B", and a person within area C will be referred to as "person C".

[0048]

After that, human detection unit 32 specifies the direction of the person, based on the location within the video of the area
25 specified in step S23 (S24). Storage unit 37 stores information indicating the location of camera 22 (e.g., the center of the top of table 40) and the capture range (i.e., the angle of view) of camera 22 in advance. Human detection unit 32 can specify which direction corresponds to the location within the video.

30 [0049]

Next, human detection unit 32 estimates the distance from sensing device 20 (i.e., camera 22) to a person, based on the size of the area specified in step S23 (S25) is. In this case, the larger the area specified in step S23 is, the shorter distance is estimated from
35 sensing device 20 (i.e., camera 22) to the person. Note that in step

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 location of the sound source as shown in FIG. 7 (the directions represented by angles in the top view). In the following
15 embodiment, 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-

dimensional coordinates (i.e., angles) corresponding to the second location (S32). If microphone array 21 and camera 22 are in different locations, the first locations converted into the two-dimensional coordinates are corrected based on the difference.

5 [0055]

Utterance amount estimation unit 34 then checks the converted first locations against the second location to detect a speaker (i.e., the location of a speaker) (S33). FIG. 9 schematically shows a result of detecting the location of the speaker. FIG. 9 is obtained by
10 superimposing the second locations (FIG. 5) on the first locations (FIG. 7).

[0056]

For example, assume that difference $\Delta\theta_1$ in the angle between the second location of sound source S1 and the first location of person
15 A is smaller than or equal to a predetermined value. At this time, utterance amount estimation unit 34 detects that sound source S1 is person A. That is, person A is detected as a speaker. For example, assume that difference $\Delta\theta_2$ in the angle between the second location of sound source S2 and the first location of person C is smaller than
20 or equal to a predetermined value. At this time, utterance amount estimation unit 34 detects that sound source S2 is person C. That is, person C is detected as a speaker.

[0057]

Utterance amount estimation unit 34 can track people A, B, and
25 C and estimate the amounts of utterance of people A, B, and C by repeating the operation in FIG. 8 at each unit time. Specifically, utterance amount estimation unit 34 can estimate the periods in which people A, B, and C are detected as a speaker, as the periods in which people A, B, and C are speaking. That is, utterance amount
30 estimation unit 34 can store information indicating the periods in which people A, B, and C are speaking (i.e., information indicating the amounts of utterance) in storage unit 37. FIG. 10 shows example information indicating the amounts of utterance. As shown in FIG. 10, the information indicating the amounts of utterance obtained

includes the ID information assigned in step S23 described above in association with the amounts of utterance (the periods of utterance).
[0058]

5 In this manner, utterance amount estimation unit 34 can track people based on the first locations detected by human detection unit 32 and the second location detected by sound source detection unit 31. Utterance amount estimation unit 34 can also estimate the amounts of utterance of people within space 70 (or the amount of utterance of a subject). Such a method of estimating the amounts
10 of utterance using utterance amount estimation unit 34 is useful to analyze the quality of communication at a conference, for example, at which people change their seats. The "change [in] their seats" here includes, for example, a case that someone moves to use whiteboard 60.

15 [0059]

Utterance amount estimation unit 34 can also estimate the amounts of utterance of people, while protecting the anonymity of the people without specifying individuals by voice or image recognition.
[0060]

20 [Operation of Head orientation Estimation Unit]

Now, an operation of head orientation detection unit 33 will be described more in detail. FIG. 11 is a flowchart of the operation of head orientation detection unit 33. FIG. 12 shows a three-dimensional coordinate space for illustrating the operation of head
25 orientation detection unit 33. As shown in FIG. 12, the following description of the operation of head orientation detection unit 33 assumes that camera 22 is placed not at center O (0, 0, 0) of table 40 but in location C (x0, y0, z0) on table 40. Note that FIG. 12 shows the coordinates in a video represented by U-V, in addition to the three-dimensional coordinates in space 70 represented by X-Y-Z.
30 [0061]

First, head orientation detection unit 33 obtains video information about a video obtained in space 70 from camera 22 of sensing device 20 (S41).

35 [0062]

Next, head orientation detection unit 33 specifies the orientation of the head of a person in the video based on the obtained video information (S42). Head orientation detection unit 33 can detect the orientation of the head of the person by specifying an area
5 corresponding to the head of the person in the video, for example, by face recognition processing, and applying a machine learning model to the specified area. As shown in FIG. 12, the vector indicating the orientation of the head at this time has an angle A° with respect to the line (i.e., the straight line) connecting location P of the person
10 and location C of the camera.
[0063]

Head orientation detection unit 33 then obtains the location of the person (more specifically, the location of the head of the person) detected by human detection unit 32 (S43). As shown in FIG. 12,
15 the person is at this time located at P (x_1, y_1, z_1).
[0064]

For example, x_1 can be expressed by the equation below using $z_1 - z_0$ determined by the actual size of table 40, horizontal angle α of view of camera 22, the location (u, v) of a person in a video, and
20 transverse size w of the video. Note that the information such as the size of table 40, horizontal angle α of view of camera 22, and transverse size w of the video is stored in advance in storage unit 37.
[0065]

$$x_1 = x_0 + ((u - w/2)/w) \times (z_1 - z_0) \times \tan(\alpha)$$

25 [0066]

After that, head orientation detection unit 33 corrects the orientation (i.e., angle A) of the head with reference to the location of camera 22 specified in step S42, based on the location of a person (specifically, coordinate x_1) obtained in step S43 (S44). FIG. 13
30 illustrates such correction of an angle (i.e., a view of table 40 from above).
[0067]

Head orientation detection unit 33 can obtain $\angle OPC$ in FIG. 13 based on coordinate x_1 obtained in step S43, and can use $A + \angle OPC$
35 as the angle after the correction.

[0068]

Head orientation detection unit 33 can track the orientation of the head of a person by repeating the operation in FIG. 11 at each unit time.

5 [0069]

[Operation of Attention Determination Unit]

Now, an operation of attention determination unit 35 will be described more in detail. FIG. 14 is a flowchart of the operation of attention determination unit 35.

10 [0070]

Attention determination unit 35 obtains the first locations detected by human detection unit 32 and the orientation of the head of a person detected by head orientation detection unit 33 (S51). The first locations obtained at this time are more specifically the first locations of people within space 70. The "orientation of the head of a person" includes the orientations of the heads of these people.

[0071]

Next, attention determination unit 35 determines the target direction of a subject who is one of the people (S52). FIG. 15 illustrates the target direction of the subject (i.e., a view of table 40 from above).

[0072]

For example, assume that the subject is the person at location P. In this case, attention determination unit 35 determines, as the target direction, the direction from location P to location G of another person. Note that the target is not necessarily a person but may be display 50 or whiteboard 60 as will be described later.

[0073]

Attention determination unit 35 then determines an allowable range around the target (S53). For example, in the example of FIG. 15, the allowable range is determined to extend from $\angle GPO - \beta$ to $\angle GPO + \beta$. B is a coefficient for determining the allowable range. Note that value β depends on which one of a person, display 50, and whiteboard 60 is the target.

35 [0074]

Next, attention determination unit 35 determines whether the subject is paying attention by comparing the orientation of the head of the subject as obtained in step S51 to the allowable range determined in step S53 (S54). The attention here means the
5 attention to communication among people including the subject in space 70. If the orientation of the head of the subject is within the allowable range, the subject is believed to see another person. Attention determination unit 35 thus determines that the subject pays attention to the communication. If the orientation of the head of the
10 subject is out of the allowable range, the subject is not believed to see another person. Attention determination unit 35 thus determines that the subject pays no attention to the communication.
[0075]

Attention determination unit 35 can determine the attention of
15 people A, B, and C by repeating the operation in FIG. 14 for each of people (i.e., people A, B, and C) as the subject at each unit time. Specifically, attention determination unit 35 can integrate the period in which each of people A, B, and C is determined to pay attention and store information indicating the periods in which the people pay
20 attention to communication in storage unit 37. FIG. 16 shows example information indicating the periods of attention. As shown in FIG. 16, the information indicating the amounts of utterance includes the ID information assigned in step S23 described above in association with the periods of attention.
25 [0076]

In this manner, attention determination unit 35 can determine whether the subject is paying attention to communication among people within space 70, based on the following information. The information includes the first locations detected by human detection
30 unit 32 and the orientation of the head detected by head orientation detection unit 33.
[0077]

Note that attention determination unit 35 may further obtain the second location detected by sound source detection unit 31 to
35 detect the location of the speaker or obtain the location of the speaker

detected by utterance amount estimation unit 34. In each case, attention determination unit 35 can determine whether the subject is paying attention, taking into consideration whether a person is a speaker. For example, only when the head of the subject is directed
5 toward the speaker, attention determination unit 35 can determine that the subject pays attention to communication. When the subject sees a person who is not speaking, attention determination unit 35 can determine that the subject pays no attention to communication.
[0078]

10 When taking into consideration whether a person is a speaker, attention determination unit 35 may determine whether the subject is paying attention as follows.
[0079]

For example, assume that there is a conference among people
15 in space 70 and person A right beside a subject (e.g., person C) is speaking. It is difficult for the subject to face the person right beside him/her. The subject may see another person B. The subject may also face display 50 (in FIG. 2) or whiteboard 60 (in FIG. 2) displaying report materials, for example, for the conference. Assume that the
20 subject is determined to pay attention to communication, only when the head of the subject is directed toward the speaker as described above. In this case, there is a problem that the subject who actually pays attention to the communication is determined to pay no attention to the communication.

25 [0080]

To address the problem, in such a case, the subject may be determined to pay attention to communication in the following periods, while person A is speaking: (1) the period in which the subject faces person A; (2) the period in which the subject faces display 50; (3)
30 the period in which the subject faces a person (person B) other than person A; and (4) the period in which the subject faces whiteboard 60. That is, attention determination unit 35 may determine the period in which the head of the subject detected faces any of a person other than the subject, display 50, and whiteboard 60, as the period
35 in which a person pays attention to communication, when one of the

people other than the subject is speaking. Note that the period in which the subject faces whiteboard 60 may be determined as a period in which the subject pays attention to communication, only when A is speaking near the whiteboard.

5 [0081]

[Operation of Communication Analysis Unit]

Now, an operation of communication analysis unit 36 will be described more in detail. FIG. 17 is a flowchart of the operation of communication analysis unit 36.

10 [0082]

Communication analysis unit 36 reads, from storage unit 37, the information indicating the amounts of utterance (FIG. 10) and information indicating the periods of attention (FIG. 16) stored in storage unit 37 (S61).

15 [0083]

Next, communication analysis unit 36 analyzes the quality of communication in space 70 based on the obtained information indicating the amounts of utterance and the obtained information indicating the periods of attention. For example, the communication analysis unit scores the quality of communication using the following scoring criteria.

20

[0084]

For example, communication analysis unit 36 calculates the ratios of the utterance of people A to C, based on the obtained information indicating the amounts of utterance. The closer to 1 the minimum/maximum of the ratios is (i.e., the more evenly people A to C are speaking), the higher the first score of the amounts of utterance are. Specifically, if the ratios of the utterance of people A to C are 1:1.2:1.5, communication analysis unit 36 obtains the first score based on the difference between 1 (i.e., the minimum)/1.5 (i.e., the maximum) and 1.

25

30

[0085]

For example, communication analysis unit 36 sets a second score as follows, based on the obtained information indicating the periods of attention. The higher the average of the periods of

35

attention of people A to C is, the higher the second score of attention is.

[0086]

Communication analysis unit 36 obtains the sum of the first
5 and second scores as the final score indicating the quality of
communication in space 70. Each of people A to C can check the final
score indicating the quality of communication (i.e., a result of
analyzing the quality of communication) by accessing communication
analysis system 10 (i.e., information processing system 30) using an
10 information terminal, such as a smartphone or a personal computer.
FIG. 18 shows an example display screen showing, on an information
terminal, a score indicating the quality of communication.

[0087]

In this manner, communication analysis unit 36 can analyze the
15 quality of communication among people within space 70, based on a
result of assuming the amounts of utterance and a result of
determining whether the subject is paying attention. Note that
communication analysis unit 36 only needs to analyze the quality of
communication among people within space 70, based on at least one
20 of a result of assuming the amounts of utterance or a result of
determining whether the subject is paying attention. The scoring
criteria described above is a mere example. The score may be
determined as appropriate depending on what is required for the
communication in space 70.

25 [0088]

[Variation]

In the embodiment described above, microphone array 21,
camera 22, and distance sensor 23 are placed on table 40.
Alternatively, microphone array 21, camera 22, and distance sensor
30 23 may be placed on a ceiling. Microphone array 21, camera 22, and
distance sensor 23 are not necessarily concentrated in one place but
may be placed in different places.

[0089]

While both sound source detection unit 31 and human detection
35 unit 32 are used to detect a speaker in the embodiment described

above, only human detection unit 32 may be used to detect a speaker. For example, human detection unit 32 can detect whether a person is speaking, based on the motion or other characteristics of the person in a video.

5 [0090]

In the embodiment described above, communication analysis unit 36 may calculate the final score indicating the quality of communication (i.e., a result of analyzing the quality of communication) in real time at a conference in space 70, and control
10 the environment in space 70 in real time based on the calculated final score.

[0091]

For example, assume that communication analysis unit 36 determines that the calculated final score is lower than a
15 predetermined value (i.e., the communication is inactive). Communication analysis unit 36 then transmits a control signal to an environment control device (not shown) placed in space 70, thereby controlling the environment in space 70.

[0092]

20 Specifically, the environment control device includes an air conditioner that controls the temperature environment in space 70, an illumination device that controls the light environment in space 70, an aroma diffuser that controls the smell in space 70, a music player that controls the sound environment in space 70.

25 [0093]

For example, communication analysis unit 36 can make the communication in space 70 active by setting a higher temperature for the air conditioner or a higher brightness than the current brightness for the illumination device. Communication analysis unit 36 may also
30 make the communication in space 70 active by activating the aroma diffuser in space 70 or causing the music player in space 70 to play music.

[0094]

[Advantageous Effects]

As described above, attention determination system 39 includes: human detection unit 32 that obtains video information about a video including people within space 70 with display 50 and whiteboard 60, and detects first locations which are locations of the people in space 70, based on the video information obtained; head orientation detection unit 33 that obtains the video information, and detects an orientation of a head of a subject among the people, based on the video information obtained; and attention determination unit 35 that determines whether the subject is paying attention to communication among the people within space 70, based on the first locations detected and the orientation of the head of the subject detected. Attention determination unit 35 determines a period in which the head of the subject detected faces any of one of the people other than the subject, display 50, and whiteboard 60, as a period in which the subject pays attention to the communication, when one of the people other than the subject is speaking.

[0095]

Such attention determination system 39 can determine how much the subject pays attention to the communication.

[0096]

For example, table 40 is placed in space 70. Human detection unit 32 obtains the video information from camera 22 on table 40.

[0097]

Such attention determination system 39 can obtain the video information from camera 22 on table 40.

[0098]

For example, a plurality of cameras 22 are placed in space 70. The video information obtained by human detection unit 32 includes video information about a video captured by each of the plurality of cameras 22.

[0099]

Such attention determination system 39 can obtain the video information from the plurality of cameras 22 on table 40.

[0100]

For example, distance sensor 23 that measures a distance from camera 22 to the subject is placed in space 70. Human detection unit 32 detects the first locations based on the video information obtained and a result of detection by distance sensor 23.

5 [0101]

Such attention determination system 39 can detect the first locations based on the result of detection by distance sensor 23.

[0102]

For example, human detection unit 32 estimates a distance
10 from human detection unit 32 to the subject, based on a size of the subject in the video, to detect the first locations.

[0103]

Such attention determination system 39 can estimate the distance from human detection unit 32 to the subject, based on the
15 size of the subject in the video.

[0104]

Communication analysis system 10 includes attention determination system 39 and communication analysis unit 36 that analyzes a quality of communication among people within space 70,
20 based on whether the subject is paying attention determined.

[0105]

Such communication analysis system 10 can analyze the quality of communication, based on whether the subject is paying attention.

[0106]

For example, communication analysis system 10 further includes sound source detection unit 31 that obtains sound information about a sound obtained in space 70, and detects a second location which is a location of a sound source in space 70, based on the sound information obtained; and utterance amount estimation
30 unit 34 that tracks the subject based on the first locations detected and the second location detected, and estimates an amount of utterance of the subject being tracked. Communication analysis unit 36 analyzes the quality of communication, based on whether the subject is paying attention determined and the amount of the
35 utterance of the subject as estimated.

[0107]

Such communication analysis system 10 can analyze the quality of communication, based on whether the subject is paying attention and the amount of utterance of the subject.

5 [0108]

An attention determination method is to be executed by a computer, such as attention determination unit 35. The attention determination method includes: obtaining video information about a video including people within space 70 with display 50 and whiteboard 60, and detecting first locations which are locations of the people in space 70, based on the video information obtained; obtaining the video information, and detecting an orientation of a head of a subject among the people, based on the video information obtained; and determining whether the subject is paying attention to communication among the people in space 70, based on the first locations detected and the orientation of the head of the subject detected. The determining whether the subject is paying attention includes determining a period in which the head of the subject detected faces any of one of the people other than the subject, display 50, and whiteboard 60, as a period in which the subject pays attention to the communication, when one of the people other than the subject is speaking.

[0109]

Such an attention determination method can determine how much the subject pays attention to the communication.

[0110]

[Other Embodiments]

While the embodiment has been described above, the present invention is not limited to the embodiment described above.

30 [0111]

For example, in the embodiment described above, the communication analysis system includes the plurality of devices but may be a single device. For example, the communication analysis system may be a single device corresponding to the information processing system, the speaker diarisation system, or the attention

determination system. If the communication analysis system includes a plurality of devices, the functional elements of the communication analysis system may be divided into the devices in any way.

5 [0112]

The communication method between the devices in the embodiment described above is not particularly limited. A relay device (not shown) may be interposed between the devices.

[0113]

10 In the embodiment described above, the processing executed by a specific processor may be executed by another processor. The plurality of processing may be executed in a different order or in parallel with each other.

[0114]

15 In the embodiment described above, the elements may be achieved by executing software programs suitable for the elements. The elements may be achieved by a program executer, such as a CPU or a processor, reading and executing a software program stored in a storage medium, such as a hard disk or a semiconductor memory.

20 [0115]

The elements may be achieved by hardware. For example, the elements may be circuits (or an integrated circuit). These circuits may form a circuit as a whole or may be separate circuits. These circuits may be general-purpose or dedicated circuits.

25 [0116]

Note that these general and specific aspects of the present invention may be implemented using a system, a device, a method, an integrated circuit, a computer program, or a computer-readable recording medium such as a CD-ROM, or any combination of systems, devices, methods, integrated circuits, computer programs, or recording media.

30

[0117]

For example, the present invention may be directed to a method of estimating the amount of utterance, to be executed by a computer, such as a speaker diarisation system. The present

35

invention may also be a program for causing a computer to execute such a method of estimating the amount of utterance, or a non-transitory computer-readable recording medium having recorded thereon such a program.

5 [0118]

The present invention may be directed to a method of determining whether a subject is paying attention, to be executed by a computer, such as an attention determination system. The present invention may also be a program for causing a computer to execute
10 such a method of determining whether the subject is paying attention, or a non-transitory computer-readable recording medium having recorded thereon such a program.

[0119]

The present invention may include forms obtained by various
15 modifications to the foregoing embodiment that can be conceived by those skilled in the art or forms achieved by freely combining the elements and functions in the foregoing embodiment without departing from the scope and spirit of the present invention.

[Reference Signs List]

20 [0120]

10	communication analysis system
20	sensing device
21	microphone array
22	camera
25	23 distance sensor (sensor)
	30 information processing system
	31 sound source detection unit
	32 human detection unit
	33 head orientation detection unit
30	34 utterance amount estimation unit
	35 attention determination unit
	36 communication analysis unit
	37 storage unit
	38 speaker diarisation system
35	39 attention determination system

40	table
50	display
60	whiteboard
70	space

[CLAIMS]

[Claim 1]

An attention determination system comprising:

5 a human detection unit that obtains video information about a video including people within a space in which a display and a whiteboard are placed, and detects first locations which are locations of the people in the space, based on the video information obtained;

10 a head orientation detection unit that obtains the video information, and detects an orientation of a head of a subject among the people, based on the video information obtained; and

an attention determination unit that determines whether the subject is paying attention to communication among the people in the space, based on the first locations detected and the orientation of the head of the subject as detected, wherein

15 when one of the people other than the subject is speaking, the attention determination unit determines, as a period in which the subject pays attention to the communication, a period in which the head of the subject detected faces any of (i) one of the people other than the subject, (ii) the display, and (iii) the whiteboard.

20

[Claim 2]

The attention determination system according to claim 1, wherein

a table is placed in the space, and

25 the human detection unit obtains the video information from a camera on the table.

[Claim 3]

30 The attention determination system according to claim 2, wherein

a plurality of cameras are placed in the space, and

the video information obtained by the human detection unit includes video information about a video captured by each of the plurality of cameras.

35

[Claim 4]

The attention determination system according to claim 2, wherein

5 a sensor that measures a distance from the camera to the subject is placed in the space, and

the human detection unit detects the first locations based on the video information obtained and a result of detection by the sensor.

[Claim 5]

10 The attention determination system according to claim 1, wherein

the human detection unit estimates a distance from the human detection unit to the subject, based on a size of the subject in the video, to detect the first locations.

15

[Claim 6]

A communication analysis system comprising:

the attention determination system according to any one of claims 1 to 5; and

20 a communication analysis unit that analyzes a quality of communication among the people in the space, based on whether the subject is paying attention determined.

[Claim 7]

25 The communication analysis system according to claim 6, further comprising:

30 a sound source detection unit that obtains sound information about a sound obtained in the space, and detects a second location which is a location of a sound source in the space, based on the sound information obtained; and

an utterance amount estimation unit that tracks the subject based on the first locations detected and the second location detected, and estimates an amount of utterance of the subject being tracked, wherein

35 the communication analysis unit analyzes the quality of the

communication, based on whether the subject is paying attention determined and the amount of the utterance of the subject as estimated.

5 [Claim 8]

An attention determination method comprising:

obtaining video information about a video including people within a space in which a display and a whiteboard are placed, and detecting first locations which are locations of the people in the space,
10 based on the video information obtained;

obtaining the video information, and detecting an orientation of a head of a subject among the people, based on the video information obtained; and

determining whether the subject is paying attention to
15 communication among the people in the space, based on the first locations detected and the orientation of the head of the subject detected, wherein

the determining whether the subject is paying attention includes determining a period in which the head of the subject
20 detected faces any of one of the people other than the subject, the display, and the whiteboard, as a period in which the subject pays attention to the communication, when one of the people other than the subject is speaking.

25 [Claim 9]

A program for causing a computer to execute the attention determination method according to claim 8.