

port that is connected with a camera 200. The input interface 930 is, in specific, a Universal Serial Bus (USB) terminal. Note that the input interface 930 may be a port that is connected with a local area network (LAN). Via the input interface 930, the crowd density calculation apparatus 100 may acquire a video stream 21 from the camera  
 5 200.

[0018] The output interface 940 is a port with which a cable of an output device, such as a display, is connected. The output interface 940 is, in specific, a USB terminal or a High Definition Multimedia Interface (HDMI, registered trademark) terminal. The display is, in specific, a liquid crystal display (LCD). Via the output interface 940, the  
 10 crowd density calculation apparatus 100 presents on the display, an analysis result outputted by the result output unit 130.

[0019] The communication device 950 communicates with other device via a network. The communication device 950 has a receiver and a transmitter. The communication device 950 is connected with a communication network, such as a LAN, Internet, or a  
 15 telephone line, through wire or by radio. The communication device 950 is, in specific, a communication chip or a network interface card (NIC). Via the communication device 950, the crowd density calculation apparatus 100 may receive the video stream 21 from the camera 200. Via the communication device 950, the crowd density calculation apparatus 100 may also transmit to an external device, the analysis result  
 20 outputted by the result output unit 130.

[0020] A crowd density calculation program is loaded into the processor 910, and executed by the processor 910. The memory 921 stores not only the crowd density calculation program, but also an operating system (OS). The processor 910 executes the crowd density calculation program while executing the OS. The crowd density  
 25 calculation program and the OS may be stored in the auxiliary storage device 922. The

The result output unit 130 outputs to an output device, such as a display, via the output interface 940, the crowd density distribution 225 outputted from the analysis unit 120.

[0026] Next, a schema of each functional component that the analysis unit 120 has will be explained.

The foreground extraction unit 121 extracts from the video frame 22, a portion that has a characteristic of foreground as a foreground image 221. Using the foreground image 221 and the analysis parameter 141 stored in the storage unit 140, the provisional density calculation unit 122 calculates provisional density distribution 222, which is apparent crowd density distribution corresponding to each position in a physical space. By determining a position where there is no person present in the physical space, the presence determination unit 123 corrects the provisional density distribution 222. The presence determination unit 123 outputs the corrected provisional density distribution 222 as corrected density distribution 223. Using a total number of persons present in an image represented by the video frame 22, the standardization unit 124 standardizes the corrected density distribution 223, and then outputs fixed density distribution 224. The distribution output unit 125 transforms the fixed density distribution 224 into an output manner, and finally outputs it as the crowd density distribution 225.

[0027] With Fig. 3, a definition of the crowd density distribution will be explained. The crowd density distribution is distribution of number of persons in a real-world physical space.

As illustrated in Fig. 3, it is assumed that, on a physical space coordinate system  $X_r$ - $Y_r$ - $Z_r$ , a surrounded region with a width of  $\Delta X$  and a depth of  $\Delta Y$  at one position  $(X_i, Y_j)$  on a plane  $X_r$ - $Y_r$  is  $\Delta S_{ij}$ , the physical space coordinate system  $X_r$ - $Y_r$ - $Z_r$

distribution 222 after implementing the presence determination process, as the corrected density distribution 223.

[0044] <<Standardization Unit>>

In step ST14, the standardization unit 124 acquires a total number of persons in the video frame 22 based on the foreground image 221. The standardization unit 124 standardizes number of persons in each three-dimensional region  $v_{ij}$  of a plurality of three-dimensional regions in the corrected density distribution 223. In specific, the standardization unit 124 standardizes crowd density by a total number of persons  $h_{total}$  present in a video frame, using Formula (1) and Formula (2) below. The standardization unit 124 calculates the total number of persons  $h_{total}$  in the video frame 22 by applying the relation 142 of the foreground area and the number of persons to the foreground image 221 overall. Rows in Formula (2) is a total number of  $i$ . And, Cols in Formula (2) is a total number of  $j$ .

[Expression 1]

$$h_{tij} = \frac{h_{total}}{h'_{total}} h_{ij} \quad (\text{Formula 1})$$

15

[Expression 2]

$$h'_{total} = \sum_{i=1}^{Rows} \sum_{j=1}^{Cols} h_{ij} \quad (\text{Formula 2})$$

coefficient  $\alpha_{lm}$  may be calculated as Formula (9), using area  $A_l$  of the three-dimensional region  $v_l$  and the duplicate region  $A_{duplm}$  of the three-dimensional region  $v_l$  and the three-dimensional region  $v_m$ , illustrated in Fig. 16.

5 [Expression 9]

$$\alpha_{lm} = \frac{A_{duplm}}{A_l} \quad (\text{Formula 9})$$

[0061] Note that it is assumed that calculation of the coefficient matrix A using Formula (9) is carried out only once in an analysis parameter load process in step ST01.

Steps ST13 and ST15 hereafter are same as those of Embodiment 1.

10 [0062] As described above, a crowd density calculation apparatus with the analysis unit 120a according to the present embodiment allows calculation of a crowd density distribution with higher accuracy by eliminating influence of a duplicate region.

[0063] Embodiment 3.

In the present embodiment, differences from Embodiment 2 will be explained.

15 Note that a component that is same as those of Embodiments 1 and 2 is denoted by a same reference sign, and description thereon may be omitted.

[0064] In Embodiment 2, a coefficient matrix A indicating duplicate relationship calculated using Formula (9) is based on an assumption that foreground appears evenly in three-dimensional region  $v_{ij}$ . Actually, the coefficient matrix A indicating the

20 duplicate relationship is influenced by a position where a person is actually present in

## CLAIMS

[Claim 1]           A crowd density calculation apparatus comprising:

a video acquisition unit to acquire a video frame from a video stream including a picked-up video image of a person; and

5                   an analysis unit to relate three-dimensional coordinates to the video frame, acquire as each three-dimensional region of a plurality of three-dimensional regions, a region representing each three-dimensional space of a plurality of three-dimensional spaces that are acquired based on the three-dimensional coordinates on the video frame, and based on number of persons present in each three-dimensional region of the

10                  plurality of three-dimensional regions, calculate density distribution of persons in the video frame as crowd density distribution,

wherein the analysis unit includes:

a foreground extraction unit to extract an image of a person in the video frame as a foreground image;

15                  a provisional density calculation unit to calculate as provisional density distribution, number of persons apparently present in each three-dimensional region of the plurality of three-dimensional regions based on the foreground image; and

a presence determination unit to determine whether there is a person present in each three-dimensional space of the plurality of three-dimensional spaces and output the

20                  provisional density distribution as corrected density distribution after number of persons in a three-dimensional region being corrected to zero, the three-dimensional region being corresponding to a three-dimensional space for which it has been determined that there is no person present.

[Claim 2]           The crowd density calculation apparatus according to claim 1,

25                  wherein each three-dimensional region of the plurality of three-dimensional

regions comprises, when the person is standing in each three-dimensional space of the plurality of three-dimensional spaces, a head region to which a head of the person corresponds and a ground region corresponding to ground on which the person is standing, and

5            wherein the presence determination unit determines, when there is a person present in both regions of the head region and the ground region in a three-dimensional region, that there is a person in a three-dimensional space corresponding to the three-dimensional region.

[Claim 3]            The crowd density calculation apparatus according to claim 1 or 2,

10            wherein the analysis unit comprises:

                 a standardization unit to acquire a total number of persons in the video frame based on the foreground image, and based on the total number of persons, standardize number of persons in each three-dimensional region of the plurality of three-dimensional regions in the corrected density distribution; and

15            a distribution output unit to acquire from the standardization unit, the standardized corrected density distribution as fixed density distribution, and transform the fixed density distribution into an output manner.

[Claim 4]            The crowd density calculation apparatus according to any one claim of claims 1 to 3,

20            wherein the analysis unit comprises a position correction unit to correct the provisional density distribution based on number of persons in a duplicate region representing a duplicate part of neighboring three-dimensional regions of the plurality of three-dimensional regions, and output the corrected provisional density distribution.

[Claim 5]            The crowd density calculation apparatus according to claim 4,

25            wherein the position correction unit repeats the correction of the provisional

density distribution until an error in the total number of persons in the video frame falls equal to or below a threshold.

[Claim 6] A crowd density calculation method comprising;

by a video acquisition unit, acquiring a video frame from a video stream

5 including a picked-up video image of a person;

by an analysis unit, relating three-dimensional coordinates to the video frame,

acquiring as each three-dimensional region of a plurality of three-dimensional regions, a region representing each three-dimensional space of a plurality of three-dimensional spaces that are acquired based on the three-dimensional coordinates on the video frame,

10 and based on number of persons present in each three-dimensional region of the plurality of three-dimensional regions, calculating crowd density distribution that is density distribution of persons in the video frame;

by a foreground extraction unit of the analysis unit, extracting an image of a person in the video frame as a foreground image;

15 by a provisional density calculation unit of the analysis unit, calculating as provisional density distribution, number of persons apparently present in each three-dimensional region of the plurality of three-dimensional regions based on the foreground image; and

by a presence determination unit of the analysis unit, determining whether  
20 there is a person present in each three-dimensional space of the plurality of three-dimensional spaces and outputting the provisional density distribution as corrected density distribution after number of persons in a three-dimensional region being corrected to zero, the three-dimensional region being corresponding to a three-dimensional space for which it has been determined that there is no person  
25 present.

[Claim 7]            A crowd density calculation program that causes a computer to execute:

                 a video acquisition process of acquiring a video frame from a video stream

5       including a picked-up video image of a person; and

                 an analysis process of relating three-dimensional coordinates to the video frame, acquiring as each three-dimensional region of a plurality of three-dimensional regions, a region representing each three-dimensional space of a plurality of three-dimensional spaces that are acquired based on the three-dimensional coordinates  
10       on the video frame, and based on number of persons present in each three-dimensional region of the plurality of three-dimensional regions, calculating crowd density distribution that is density distribution of persons in the video frame,

                 wherein the analysis process includes:

                 a foreground extraction process of extracting an image of a person in the video  
15       frame as a foreground image;

                 a provisional density calculation process of calculating as provisional density distribution, number of persons apparently present in each three-dimensional region of the plurality of three-dimensional regions based on the foreground image; and

                 a presence determination process of determining whether there is a person  
20       present in each three-dimensional space of the plurality of three-dimensional spaces and outputting the provisional density distribution as corrected density distribution after number of persons in a three-dimensional region being corrected to zero, the three-dimensional region being corresponding to a three-dimensional space for which it has been determined that there is no person present.