平成 29 年度博士学位論文

Detection of facial feature points in three-dimensional space for meal support equipment

Bo Peng

群馬大学大学院　　理工学府

理工学専攻　知能機械創製理工学領域

Contents

Chapter 1 Introduction

1.1 **The background of this research**

Currently, the number of elderly people and persons with disabilities in Japan is rapidly increasing and it becomes a big social problem. According to the 2011 edition "White Paper on Disabled Persons", the statistical result release that number of people with disabilities are 3.663 million (29 people per thousand people)[1]. On the other hand, the elderly population aged 65 and older is estimated at 29.01 million according to the survey in 2010, accounting for 22.7% of the total population [2]. In 2005, the population of the elderly was 24.88 million, 19.5% of the total population. We can understand that the elderly population was increased in a surprising rate. And it is expected that in the future as the year of 2020, the population of the elderly people will become 35.89 million. In the progress of the declining birthrate and the aging of the population, the condition of shortage of nurse/carer is getting worse, and the way to solve this situation becomes an important research issue in the future.Meanwhile, the meal, which is one of the enjoyments of the care recipient, largely depends on the care giver. Among them, there is a strong demand to eat the meals as expected, such as the pace and the order of eating, and in many nursing care facilities, a small number of carers are doing many meal support. Against this backdrop, development of a meal support manipulator has been carried out, the research of combination of industrial miniature robot and voice recognition [3], using pneumatic actuator [4] has been advanced.

1.2 **Existing products**

Some products have already been put into practical use, and SECOM Corp. sells "My spoon"[5]. In "My spoon", they prepared a manual mode for meticulous operation by using a joystick, and automatic mode for selecting a tray by pressing a button at the right time. These meal support devices have been developed in pursuing convenience, and they are very excellent in side of functionality. However, since the joystick or the button is operated by the jaw, mouth, etc. the operability of the apparatus tends to be greatly influenced by the proficiency of the user. Also, the main role is to carry food to the mouth and finally the action of putting food into the oral cavity is not a device, but the user needs to move their face by themselves, so it is necessary to develop a system which can carry food into mouth by the device entirely. And the device "Obi" which is made by Desin Corp is developed [6], and this device can catch food to the mouth by easy control, so the operability is convenient, the system set the position where the spoon reach every time, to make the patient comfortable, but the position setting should be done every time, if the position of the patient is changed even a little, this system can't reach the right position, and device should be set again, so it is better for the patient if the position of mouth is able to be detected automatically. And the Handy1 made by Rehab Robotics also can't detect the right positions [7] [8] [9]. The iARM which is made by Exact Dynamics get a good manipulator to carry food [10], but it also need a joystick to control the system, so it is hard for the patient especially hands disability, so the convenient of the control is also important.

1.3 **Face detection**

In recent years, the automatic face detection technology [11] has been developed and it has been used on home optoelectronics device, such as digital video recorder, camera, or smartphone, as a matter of course. Such recognition technique has been remarkably developed since the Viola-Jones method [12] [13] [14] [15] [16] [17] and derivative method [18] [19] had been proposed. This method can find a human face from the obtained image very easily. However, it had been constructed on the assumption that a front-facing is roughly kept. Therefore, the side face and large target cannot be recognized, and this method to detect human face is using Haar-like method [20] [21] [22], this method can also detect the human face correctly, but this system has to study about the human face feature point, so the computational complexity is very huge, in this case, it leads to the result that this method needs to study for times and hard to detect the face in real time [23] [24] [25] [26], And many other methods of face detection also have the same problem [27] [28] [29] [30].To solve this problem, the author has proposed a new method of face detection, based on facial feature color [31]. As a same approach, improvements in Viola Jones algorithm using both skin [32] [33] [34] [35] [36] [37] and eyes colors has been proposed to detect the tilted face detections [38]. Here, we used only rough skin color information to detect the nostril position. The proposed system could narrow down the candidate of the nostril by checking the color, area, and aspect ratio, the detection of nostril is proved to be stable of the human face, and according to the relative position of mouth and nostril, we narrowed down the area of mouth detection, and the accuracy rate of the mouth detection has been risen up, this mouth detection method is better than the other methods which just scan the mouth without the other areas of human [39] [40] [41] [42]. After the detection of mouth, we purpose to detect the condition of mouth by setting a threshold, according to this threshold, we can judge that whether the mouth is opened or closed, if the mouth is closed, system will not display, and manipulator will not do any movement, if the mouth is opened, the system will mark the position of mouth in real time.

1.4 **Manipulator system**

In our system, we developed a 3-link manipulator with 3 RC servomotors [43], and the control board (KONDO RCB 4 - HV: M16C / 26 A manufactured by Renesas Technology Corporation) [44], and by calculation every angle of joints, we can sent the spoon which is carrying food to the right position [45] [46] [47] [48], and when the system detected that the mouth is closed, the manipulator will be back to the inceptive position of the system.

1.5 **Summary**

Therefore, in this study, we developed a three-dimensional position measurement system of facial feature points [49] by stereo camera as a basis for constructing a meal support system that can move a spoon to the oral cavity by accurate position detection. By processing the input image obtained by the stereo camera, and detecting the face automatically, we constructed a system that detects the nostrils as facial feature points and accurately measures the mouth position in the three-dimensional space, and its effectiveness had been revealed.

The structure of this paper is shown below.

Chapter 2: The summary of this system and the description of three-dimensional measurement by stereo camera.

Chapter 3: The description of the Algorithms automatic detection of the nostrils and the mouth.

Chapter 4: The description of the control of manipulator by using three dimensional measurement systems.

Chapter 5: The summary of this paper.

Chapter 2 Construction of three-dimensional measurement system

2.1 **The construction of the system**

Fig.2.1 shows the processing flow and configuration of this system. A stereo camera consisting of two USB cameras is connected to the computer and converts the captured image into a bitmap image in real time [50]. Next, after detecting the face of the user by the image processing, the nostrils are detected as the face feature points, and the position measurement in the three-dimensional space of the object is realized by using the difference in the coordinates between the two images [51].

The overall image of the stereo camera is as shown in Fig.2.2, and the two cameras are fixed on an aluminum plate of 300 × 100 × 3 mm. set the axis in the direction of the arrow shown in Fig.2.2 with the midpoint of the camera fixed position as the origin.

In this experiment, we use two HD Webcam C615 of Logicool as USB cameras. The photograph is shown in Fig.2.3. In this camera, the CMOS sensor of 2.1 million pixels is used as an image sensor, and the focus is autofocus in 7cm-∞. The resolution can be up to 1920 × 1080 pixels (full HD 1080p) and the frame rate can be up to 30 fps, but in this research we use 640 × 480 pixel, 30 fps, RGB 24 shooting.

We used Microsoft Visual C# 2010 to construct the control program and developed GUI (Graphical User Interface) and system control as shown in Fig. 2.4 using C# as a programming language.

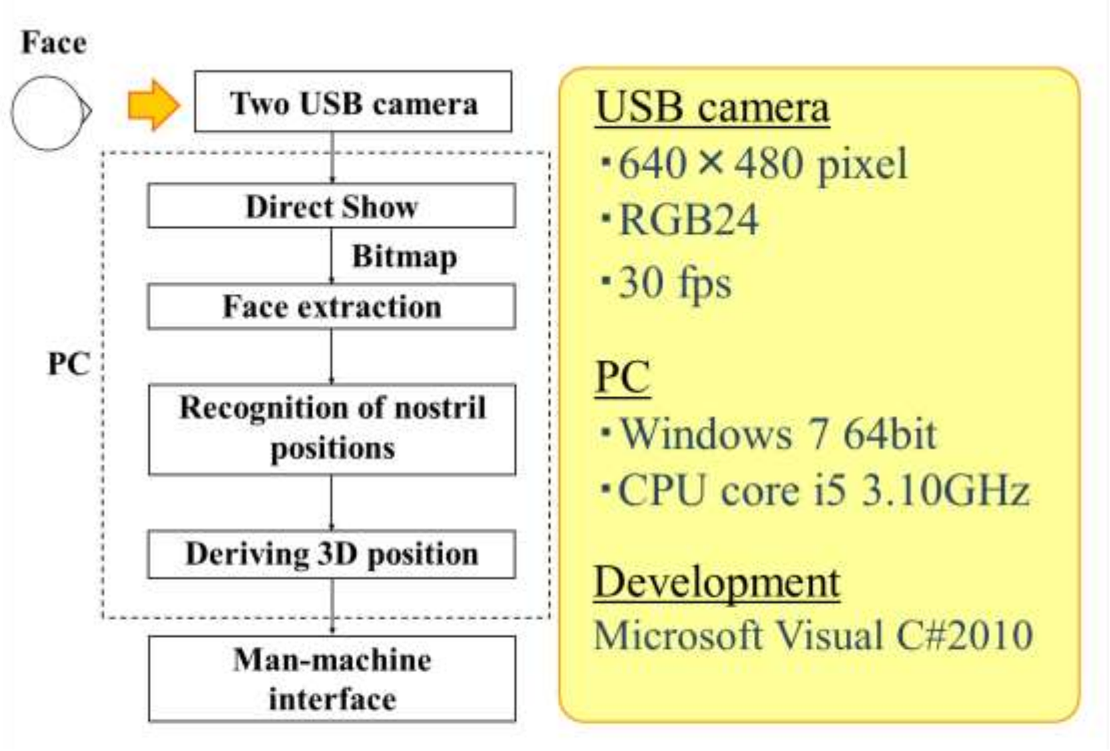


Fig. 2.1 Device system configuration

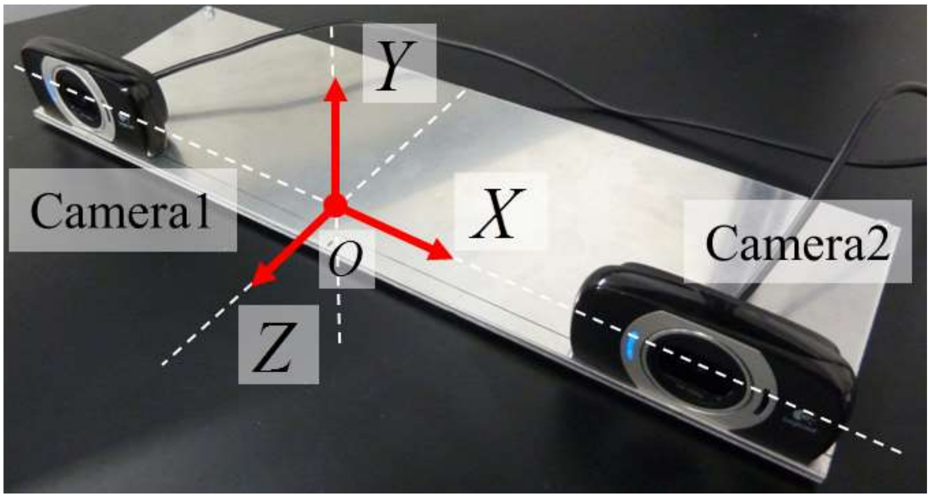


Fig.2.2　Stereo camera system

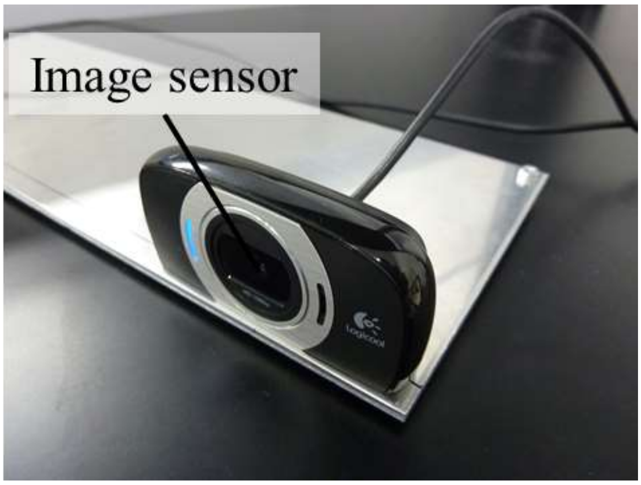


Fig. 2.3 USB camera (HD Webcam C615)

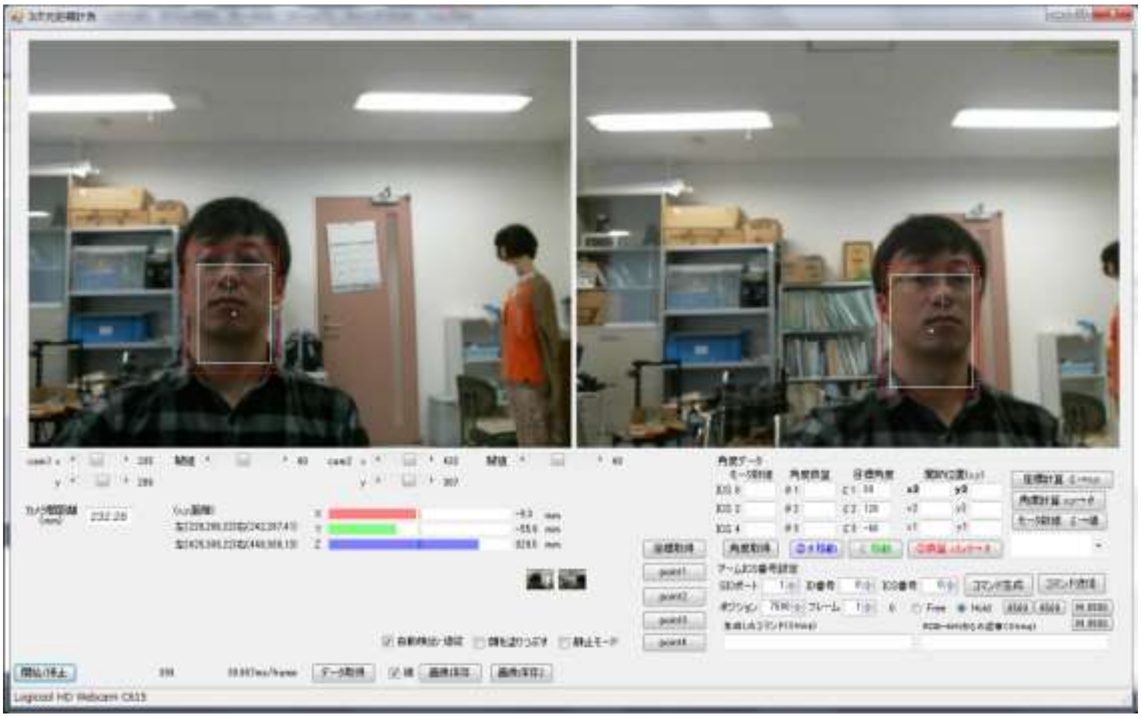


Fig. 2.4 GUI tool

2.2 **Basic formula for three-dimensional measurement by stereo camera**

In this system, the position of an object in three-dimensional space is measured by using the difference (the parallax) of the coordinate position of the measurement object displayed in the image input simultaneously from the two USB cameras. For that purpose, it is necessary to determine the angle to the object in the three-dimensional space whose origin is the camera from the coordinates on the input image first. The concrete processing will be described below.

When the coordinates (x, y) in the input image are set as shown in Fig. 2.5 and the angles θ, ϕ formed by the positive direction of the axis *X* and the axis *Y* with respect to the object are respectively as follows, the values are obtained from Fig.2.6 (a) and (b)

$$\theta = \left(\theta_0 + \frac{\alpha}{2}\right) - \frac{x}{x_{\max}} \times \alpha \tag{2.1}$$

$$\phi = \left(\phi_0 - \frac{\beta}{2}\right) + \frac{y}{y_{\max}} \times \beta \tag{2.2}$$

Here: angle θ_0 : the angle between the center axis of the camera and the axis *X*, angle ϕ_0 : the angle formed by the center axis of the camera and the axis *Y*, α : viewing angle of the camera in the axis *X* direction, β: viewing angle of the camera in the axis *Y*, $x_{\max}$: axis *X* resolution of the camera : $y_{\max}$: axis *Y* resolution of the camera, (x, y): Axial coordinate resolution of the object on the image.

Since the parameters other than x and y on the right side of the formulas (2.1) and (2.2) are constants for each system, if the values of the coordinates (x, y) acquired from the input image are substituted, the angle θ, ϕ to the object can be obtained from these equations.

Furthermore, as shown in Fig. 2.7 (a), if the distance between the cameras is *L* and the difference between the coordinates of the camera 1 and the axis *X* of the object is *l*, the following two equations are obtained

$$\tan\theta_1 = -\frac{Z}{l} \tag{2.3}$$

$$\tan\theta_2 = \frac{Z}{L - l} \tag{2.4}$$

θ_1, θ_2 are the angles obtained by substituting Equation (2.1) for Camera1 and Camera2 respectively.

From this, by rearranging the formulas (2.3) and (2.4), the coordinates to the target with respect to the axis *Z*,

$$Z = \frac{\tan\theta_1 \tan\theta_2}{\tan\theta_1 - \tan\theta_2} L \tag{2.5}$$

Furthermore, as shown in Fig. 2.7 (b) and (c), since *Z* is obtained by the formula (2.5), the value of the coordinates *X*, *Y* can be obtained and it is calculated by the following formula.

$$\begin{aligned} X &= \frac{Z}{\tan(180^\circ - \theta_1)} - \frac{L}{2} \\ &= -\frac{Z}{\tan\theta_1} - \frac{L}{2} \\ &= \left(-\frac{\tan\theta_2}{\tan\theta_1 - \tan\theta_2} - \frac{1}{2} \right) L \\ &= \frac{\tan\theta_1 + \tan\theta_2}{2(\tan\theta_2 - \tan\theta_1)} L \end{aligned} \tag{2.6}$$

$$\begin{aligned} Y &= \frac{1}{\tan\phi_1} \times \frac{\mathrm{Z}}{\sin(180^\circ - \theta_1)} \\ &= \frac{1}{\tan\phi_1} \times \frac{\mathrm{Z}}{\sin\theta_1} \\ &= \frac{1}{\tan\phi_1 \sin\theta_1} \times \frac{\tan\theta_1 \tan\theta_2}{\tan\theta_1 - \tan\theta_2} L \\ &= \frac{\tan\theta_2}{\tan\phi_1 \cos\theta_1 (\tan\theta_1 - \tan\theta_2)} L \end{aligned} \tag{2.7}$$

Therefore, by using the equations (2.5), (2.6), and (2.7), it is possible to uniquely determine the coordinates (X, Y, Z) in the three-dimensional space of the target object based on the angles θ, ϕ calculated by coordiates $(x_1, y_1), (x_2, y_2)$ which is obtained from the input images of two cameras of the system by this equation

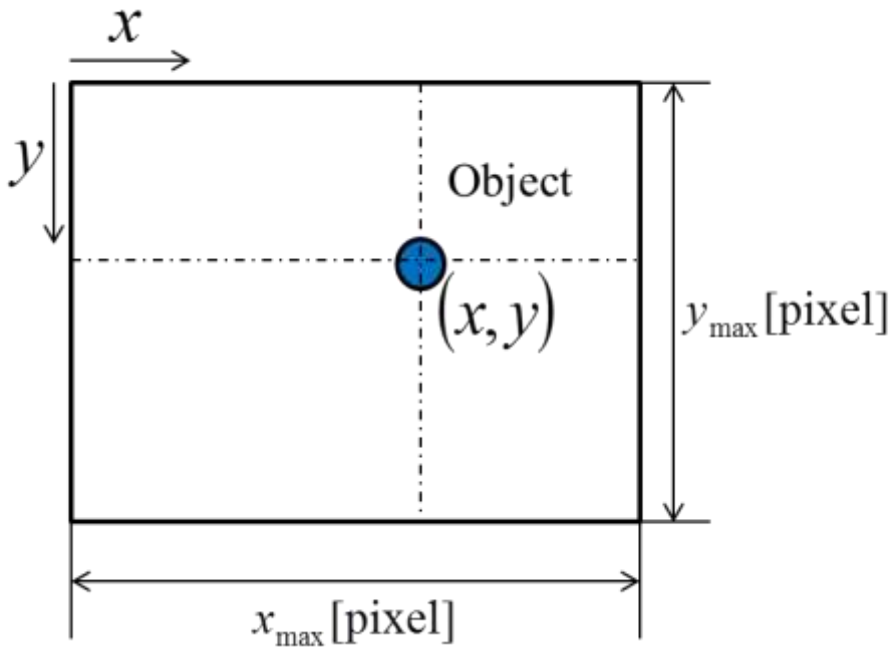
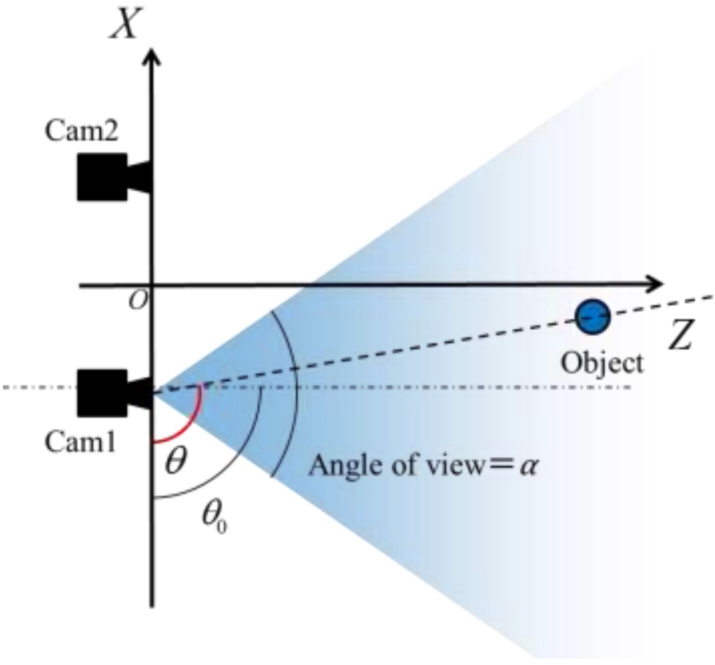
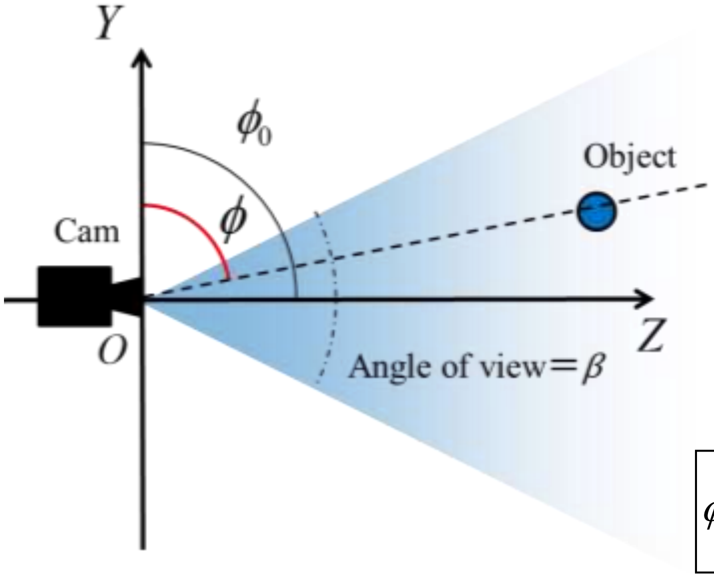


Fig.2.5 *x* and *y*-coordinate of capture image



$$\boxed{\theta = \left(\theta_0 + \frac{\alpha}{2}\right) - \frac{x}{x_{\max}} \times \alpha}$$

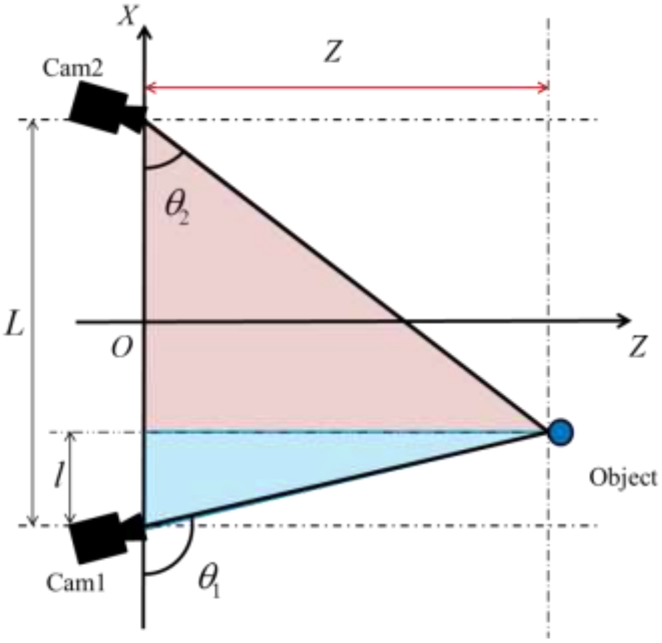
(a) X-Z plane



$$\boxed{\phi = \left(\phi_0 - \frac{\beta}{2}\right) + \frac{y}{y_{\max}} \times \beta}$$

(b) Y-Z plane

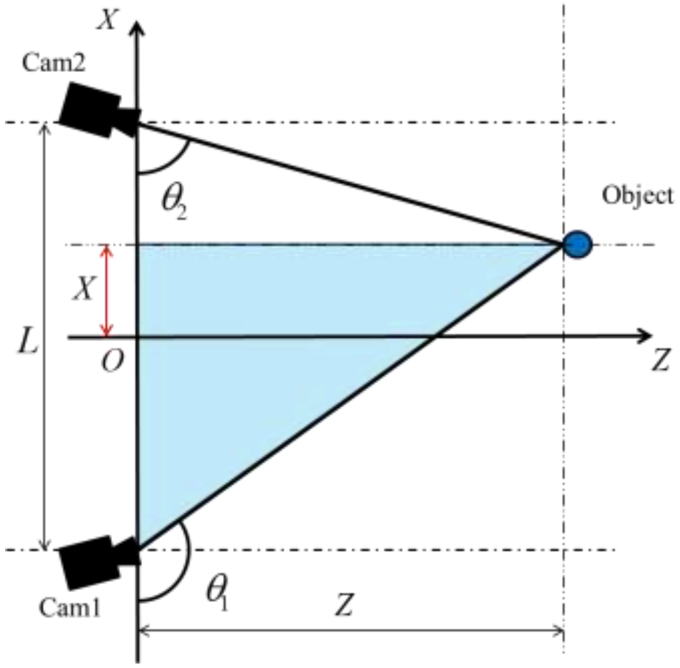
Fig. 2.6 View angle



$$\begin{cases} \tan\theta_1 = -\dfrac{Z}{l} \\ \tan\theta_2 = \dfrac{Z}{L-l} \end{cases}$$

$$\boxed{Z = \frac{\tan\theta_1 \tan\theta_2}{\tan\theta_1 - \tan\theta_2} L}$$

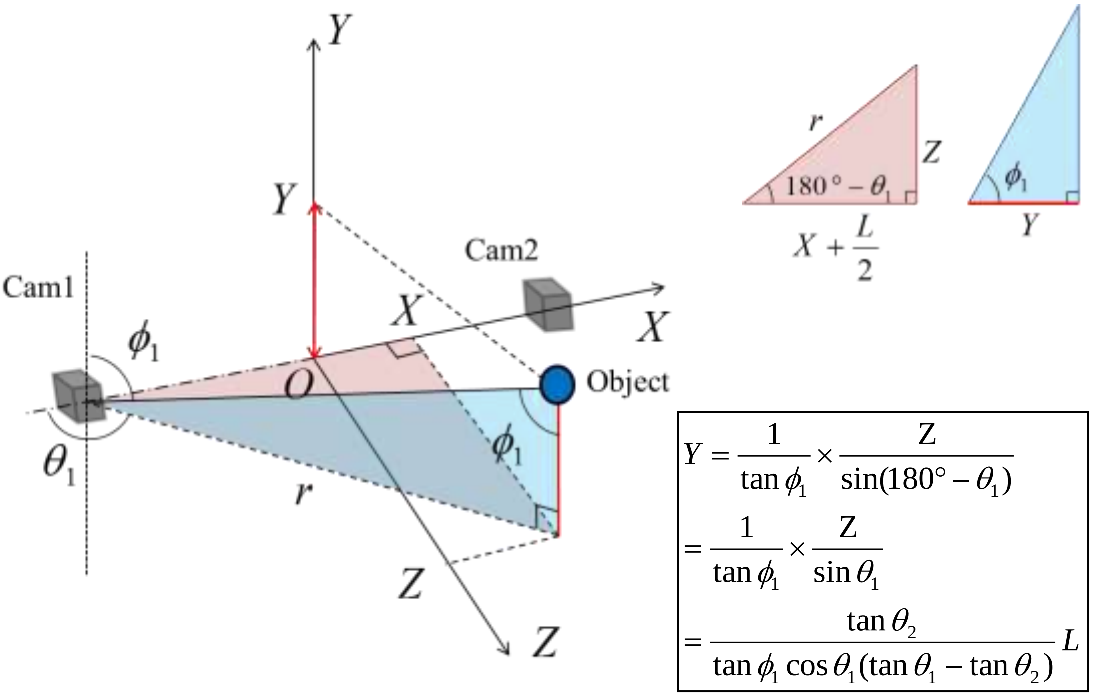
(a) Z-coordinates



$$\tan\theta_1 = -\tan(180° - \theta_1)$$
$$= -\frac{Z}{\frac{L}{2} + X}$$

$$\boxed{\begin{aligned} X &= -\frac{Z}{\tan\theta_1} - \frac{L}{2} \\ &= \frac{\tan\theta_1 + \tan\theta_2}{2(\tan\theta_2 - \tan\theta_1)} L \end{aligned}}$$

(b) X-coordinates



$$
\begin{aligned}
Y &= \frac{1}{\tan\phi_1} \times \frac{\mathrm{Z}}{\sin(180° - \theta_1)} \\
&= \frac{1}{\tan\phi_1} \times \frac{\mathrm{Z}}{\sin\theta_1} \\
&= \frac{\tan\theta_2}{\tan\phi_1 \cos\theta_1 (\tan\theta_1 - \tan\theta_2)} L
\end{aligned}
$$

(c) Y-coordinates

Fig. 2.7 Formula for computation

2.3 **Camera characteristics**

2.3.1 **Lens distortion**

The most ideal camera model for monocular cameras is the "pinhole camera model" as shown in Fig.2.8. In this model, assuming the coordinates (X, Y, Z) of the three-dimensional space and the coordinates (u, v) on the screen as the focal length f of the lens, It is known that the relationship

$$u = f \times \frac{X}{Z} \tag{2.8}$$

$$v = f \times \frac{Y}{Z} \tag{2.9}$$

is established [50].

The pinhole camera model is a model without lens distortion, and the image input from this camera becomes an ideal input without distortion as shown in Fig.2.9 (a).

However, generally wide-angle cameras have barrel type aberrations like Fig.2.9 (b), telephoto cameras have pincushion type aberrations like Fig.2.9 (c). When the input image of the camera is distorted, an accurate value cannot be obtained from the target coordinates, so it is necessary to perform calibration to correct the distortion of the lens.

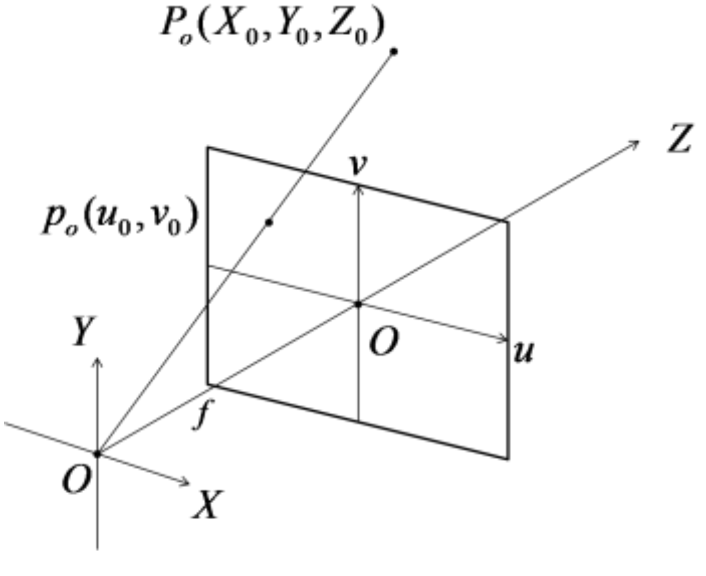
Therefore, as a preliminary experiment, when a grid-like image was taken with a resolution of 960 × 720 [pixels], it became as shown in Fig.2.10. The coordinates of the intersection point of this lattice are obtained and the scatter diagram is shown in Fig. 2.11. In the vicinity of the center, the intersection line which was almost aligned draws a gentle curve at the edge of the image, which shows that it is the same barrel type aberration as Fig.2.9 (b). From the above results, in the present system, when converting the acquired coordinates into angles and calculating the three-dimensional position, the numerical value is corrected.

2.3.2 **Camera and system parameters**

When actually performing three-dimensional measurement from the coordinates of the input image, parameters of the camera and the system such as the viewing angle α in the axial X direction, the viewing angle β in the axial Y direction, the distance L between the cameras shown in Fig.2.6 and Fig.2.7 are necessary. As a result of the verification, we decided to give the parameter $\alpha = 51.48[°]$, $\beta = 39.30[°]$, $L = 232.28[\mathrm{mm}]$ in this system.

Furthermore, in this system, the fixed angle θ_0 of the camera in the axial X direction was set as the angle correction. When imaging is performed at a resolution of 640 × 480 pixels using a camera of $\alpha = 51.48[°]$. In order to construct a precise measurement system, fixation with accuracy less than the error of $\pm 0.1°$ is required, but since processing and attachment with this accuracy is difficult, numerical correction is added to the acquired coordinates.

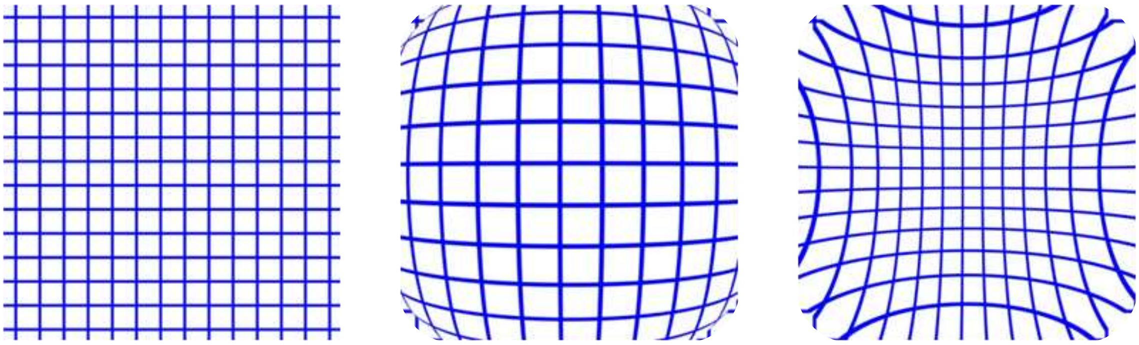
The fixed angles $\theta_{01} = 90.49[°]$, $\theta_{02} = 90.57[°]$ of cameras 1 and 2 were judged to be optimal, respectively. These parameters are used for subsequent measurement.



$$u_0 = f \times \frac{X_0}{Z_0}$$

$$v_0 = f \times \frac{Y_0}{Z_0}$$

Fig. 2.8 Pinhole camera model



(a) No distortion (b) Barrel distortion (c) Pincushion distortion

Fig. 2.9 Radial distortion

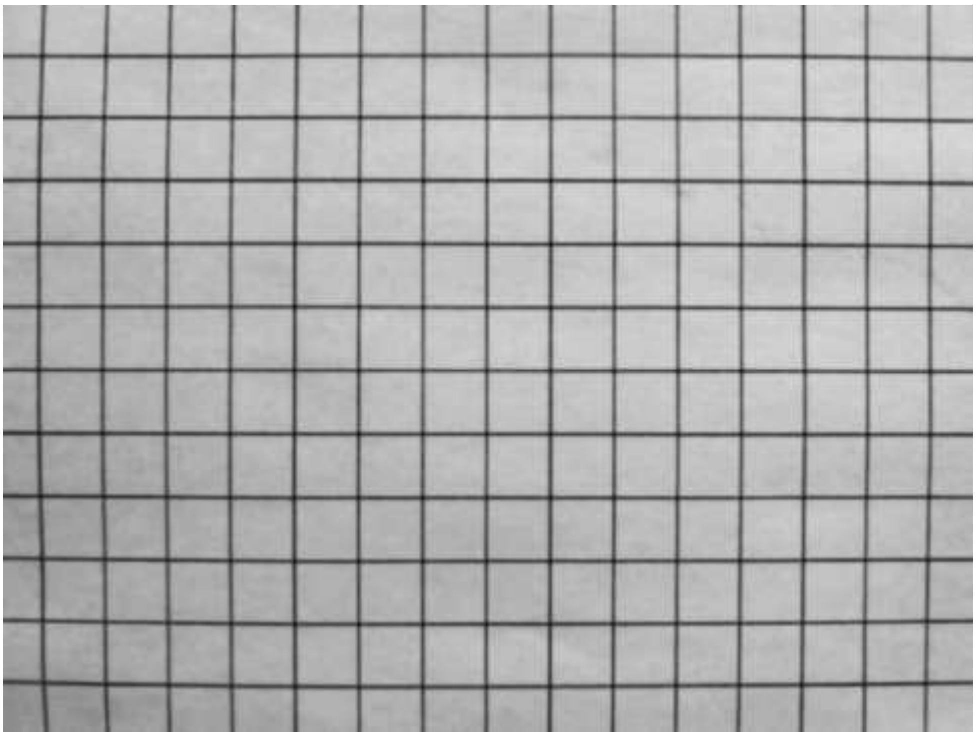


Fig. 2.10 Capture image

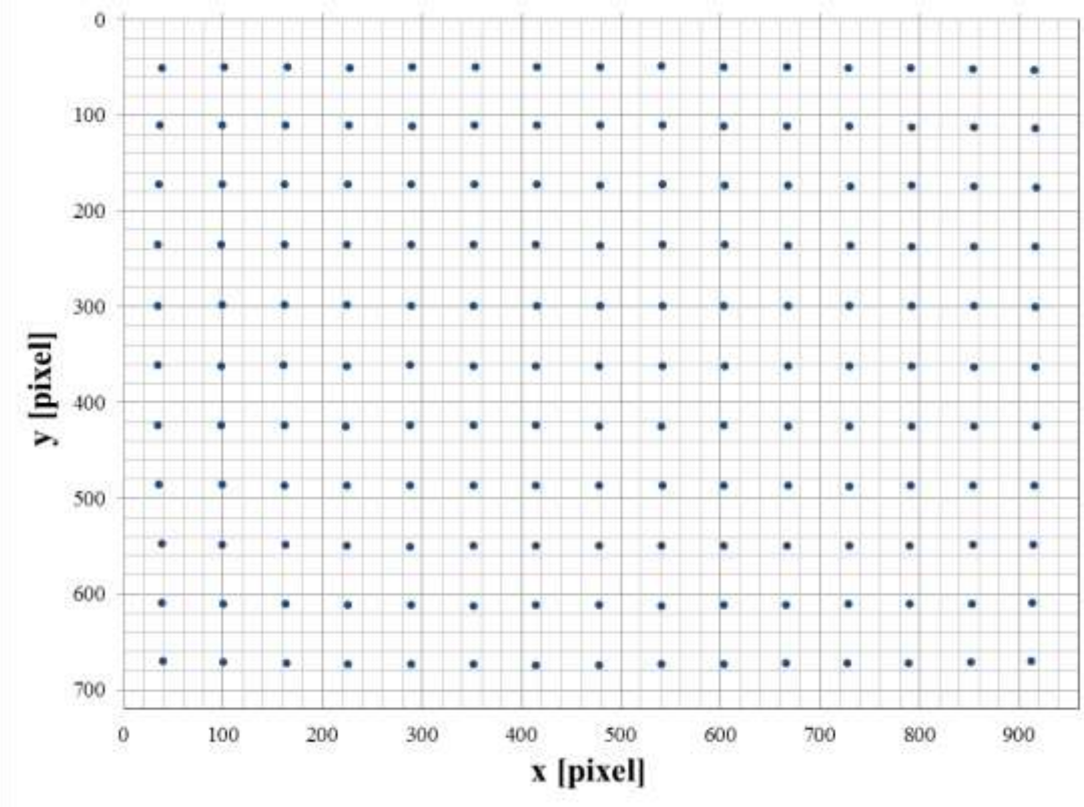


Fig. 2.11 Intersection point in capture image

2.4 **Verification experiment**

2.4.1 **Experimental method**

In this system, a verification experiment will be conducted on the error between the actual position and the theoretical value of X ,Y , Z calculated from the equations (2.5), (2.6), and (2.7) shown in Section 2.2.

As shown in Fig.2.12, the camera is fixed so that the plane $X - Y$ and the wall surface are parallel, a lattice-like image is drawn by four horizontal lines and seven vertical lines at intervals of 40 [mm] as shown in Fig.2.13Zcx t6. The coordinates x and coordinates y of each intersection in two input images are acquired, and the position in the three-dimensional space of the measurement target is obtained from the parallax.

For the experiment, the distance Z from the origin to the object was set at five points of 400 to 800 [mm] at intervals of 100 [mm].

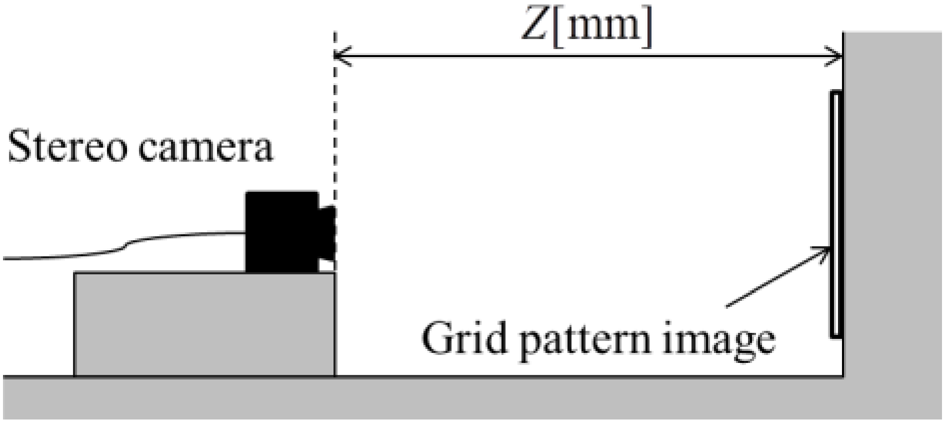


Fig. 2.12 Experimental condition

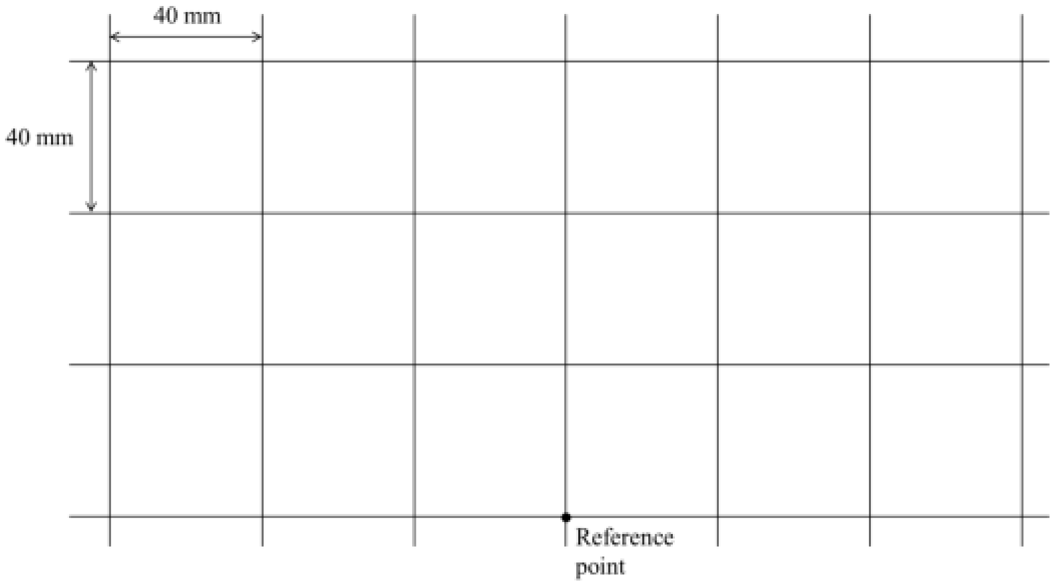


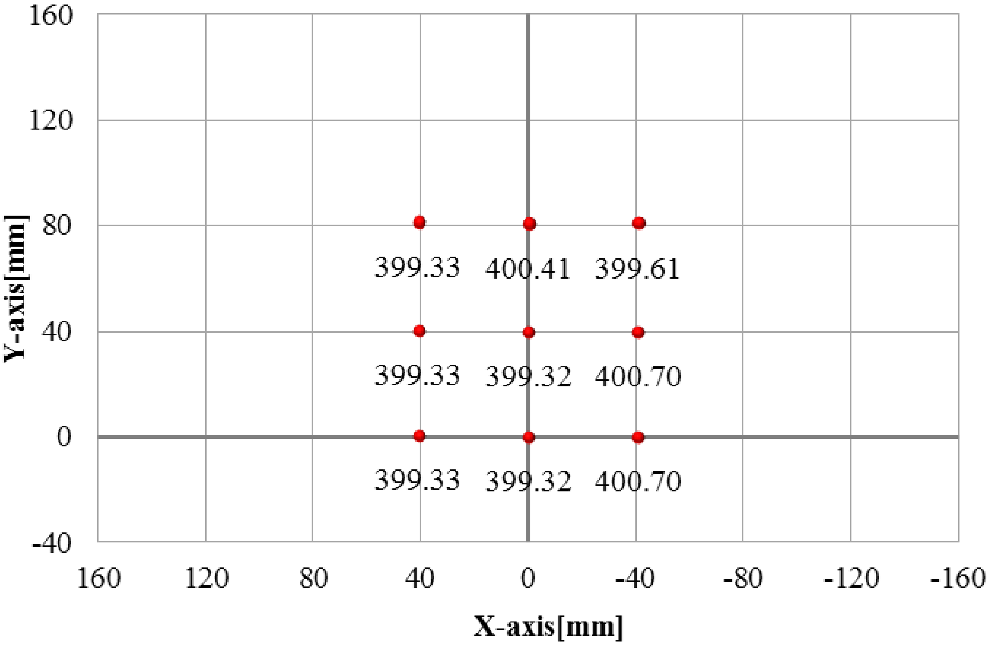
Fig. 2.13 Grid pattern image

2.4.2 Experimental results

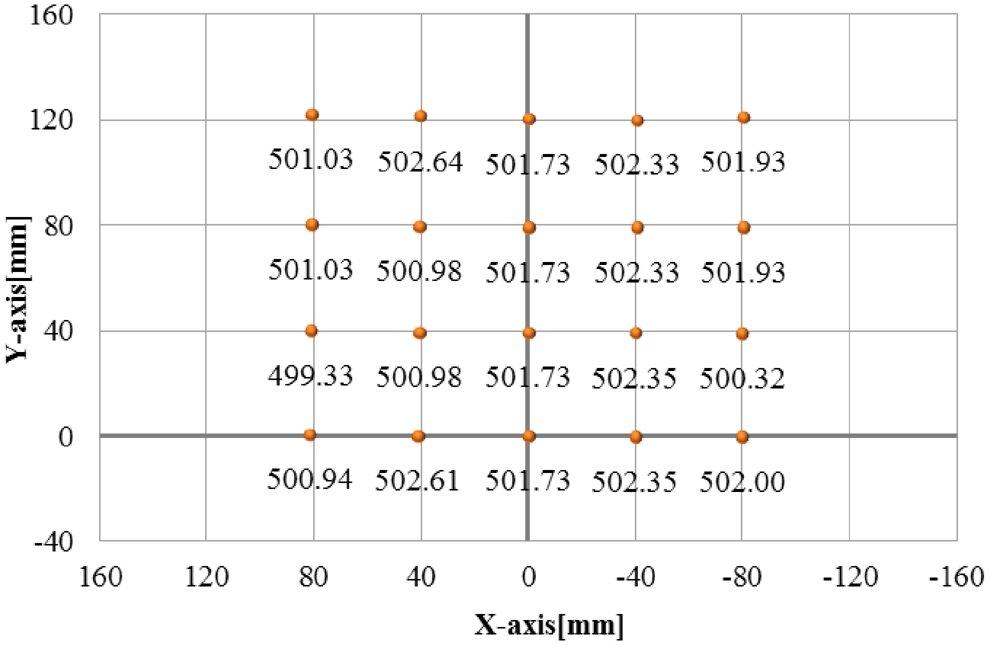
Fig.2.14 (a) - (e) shows the measurement results at each distance. Each data in the figure shows the relative position of each intersection point X , Y when the intersection point of the lowest stage in Fig.2.13 is set as the reference point $(0, 0)$ in the plane $X - Y$. Also, the values Z at each point are shown in the figure. In addition, the number of measurement results decreases for $Z = 400[\mathrm{mm}]$ and $Z = 500[\mathrm{mm}]$, but this is due to the limitation of the field of view angle of the camera.

Fig.2.15 (a) - (c) shows the distribution of error for each axis at each distance for these data. As shown in this figure, the position of the intersection point can be measured at approximately 40 [mm] intervals vertically and horizontally at any distance, It is within an error of about 0.5[%] of *X* and *Y*, and 1[%] of *Z*. And in these errors, the biggest errors is about 5[mm], by comparing with spoon size (26[mm]) and mouth size 50[mm][53], this size of error could be accepted.

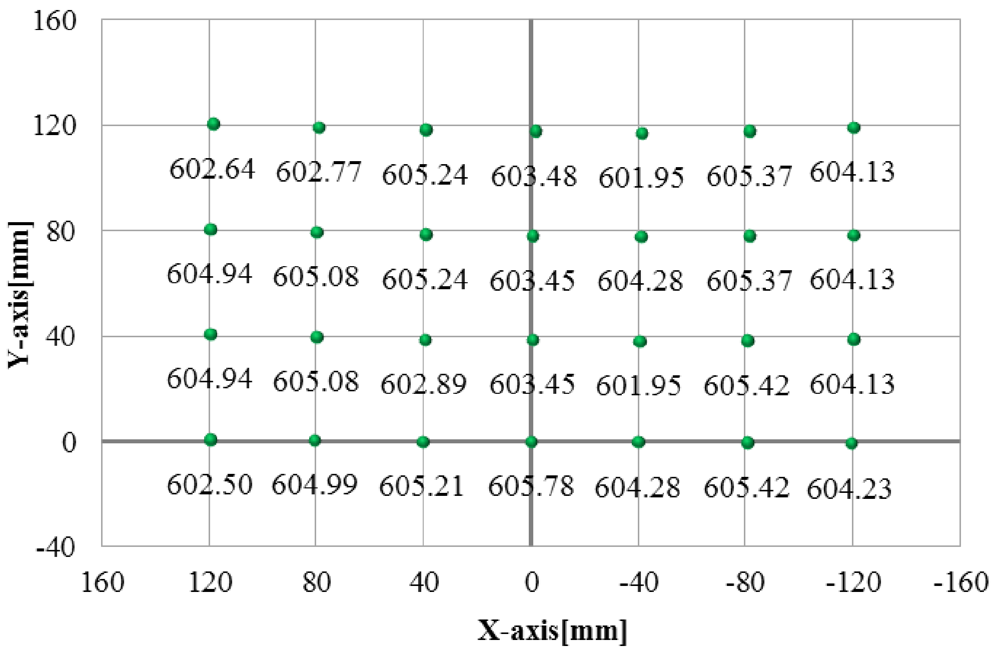
From the above, the usefulness of this system to measure the position in three dimensional spaces from the coordinates of the input image of the stereo camera was confirmed.



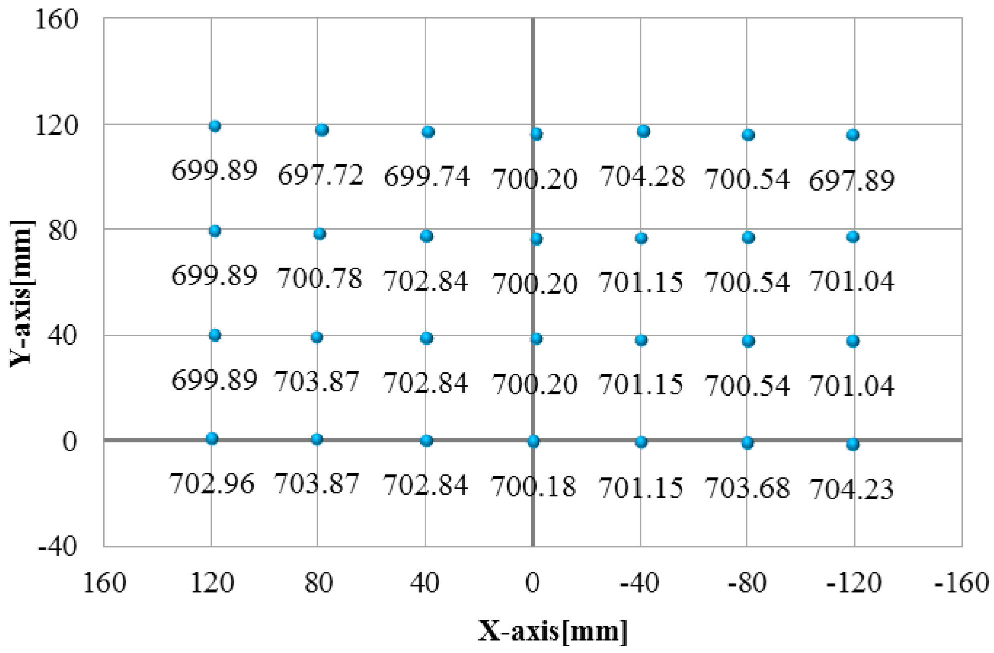
(a) $Z = 400[\text{mm}]$



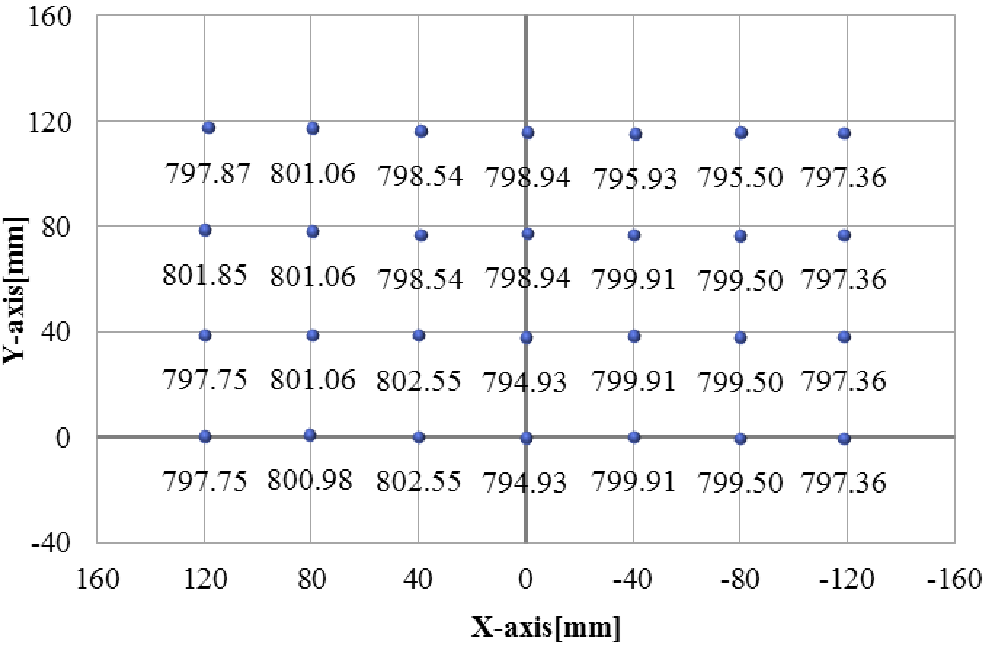
(b) $Z = 500[\text{mm}]$



(c) $Z = 600$[mm]

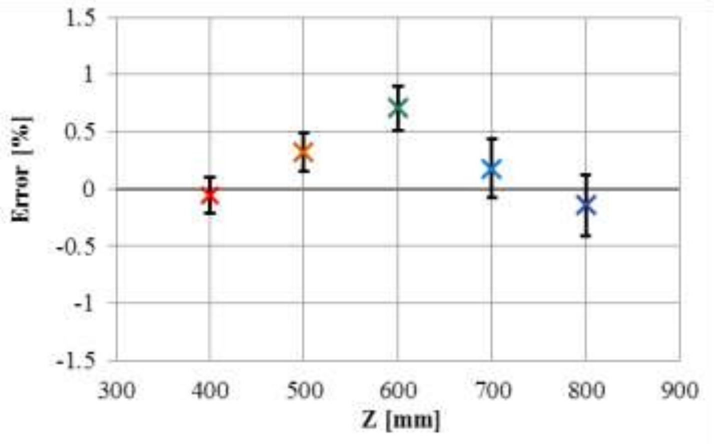


(d) $Z = 700$[mm]

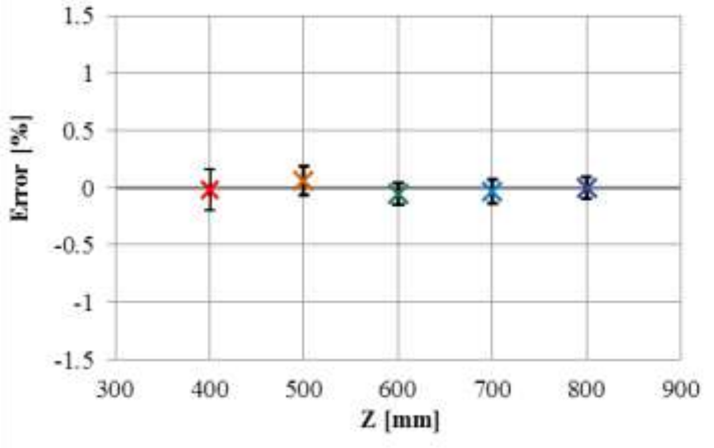


(e) $Z = 800[\mathrm{mm}]$

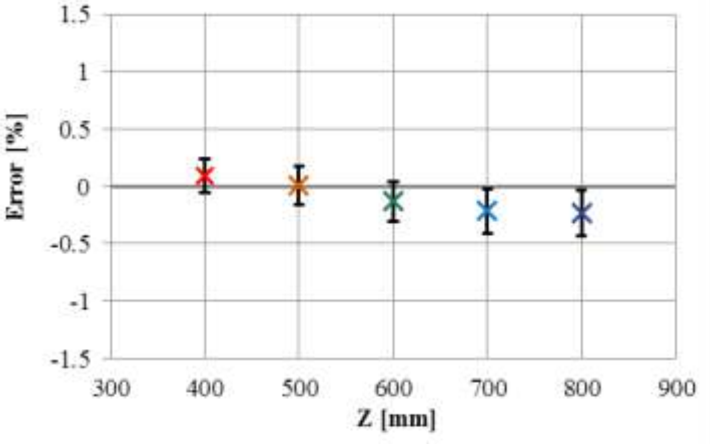
Fig. 2.14 Relative position in the *X*-*Y* plane and estimated value of *Z*



(a) *Z*-axis



(b) *X*-axis



(c) *Y*-axis

Fig. 2.15 Error in each axis

Chapter 3 Feature point detection

3.1 **Basics of image processing**

3.1.1 Binarization processing - Labeling processing

In this system, by binarizing the input image from two USB cameras, labeling processing is performed to assign numbers to each connected component and by examining the shape features such as size and center of gravity for each component, and detects nostrils as facial feature points.

For binarization, RGB luminance signals at each coordinate of the input image are used. Here, an example of binarization using the R signal will be described. Conditions for binarization are as follows.

$$g[i,j]=\begin{cases}1(R[i,j]<t)\\0(R[i,j]\geq t)\end{cases} \tag{3.1}$$

However, *g* is 1 or 0 sgn, $R[i,j]$ is the luminance of the *R* signal at the coordinates (i,j), *t* is the threshold.

As shown in Fig.3.1, a process (labeling process) is performed to add a unique label for each connected component to each pixel in the binarized image,

$$l[i,j]=\begin{cases}n(g[i,j]=1)\\0(g[i,j]=0)\end{cases} \tag{3.2}$$

n is a natural number representing the label number of the connected component to which the coordinate (i,j) belongs to.

For the image in Fig.3.2 (a), the example of the image binarized by the formula (3.1) is shown in Fig.3.2 (b), and the result of the labeling process by the equation (3.2) is shown in Fig.3.2 (c).

3.1.2 Features of connected components

After performing binarization and labeling processing, obtain the area, center of gravity, aspect ratio of the connected component. Assuming that the area S_n of the connected component of the label number n is the coordinates of the center of gravity (G_{xn}, G_{yn}),

$$S_n = \sum_i \sum_j f_n[i, j] \tag{3.3}$$

$$G_{xn} = \frac{1}{S_n} \sum_i \sum_j (f_n[i, j] \times i) \tag{3.4}$$

$$G_{yn} = \frac{1}{S_n} \sum_i \sum_j (f_n[i, j] \times j) \tag{3.5}$$

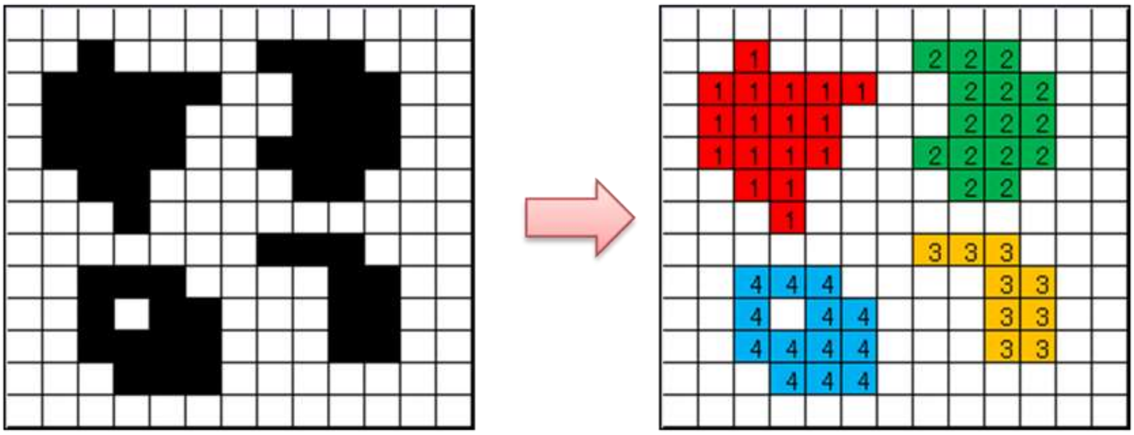
where

$$f_n[i, j] = \begin{cases} 1(l[i, j] = n) \\ 0(l[i, j] \neq n) \end{cases} \tag{3.6}$$

This is a function for determining whether or not the label number n in the coordinates (i, j) matches.

Next, the ratio (aspect ratio) between the vertical width and the horizontal width of the connected component is obtained. As shown in Fig.3.3, if the coordinates of the left end, the right end, the top end, and the bottom end of the connected component of the label number n are represented by x_{Ln}, x_{Rn}, y_{Tn}, y_{Bn}, and the aspect ratio as follows.

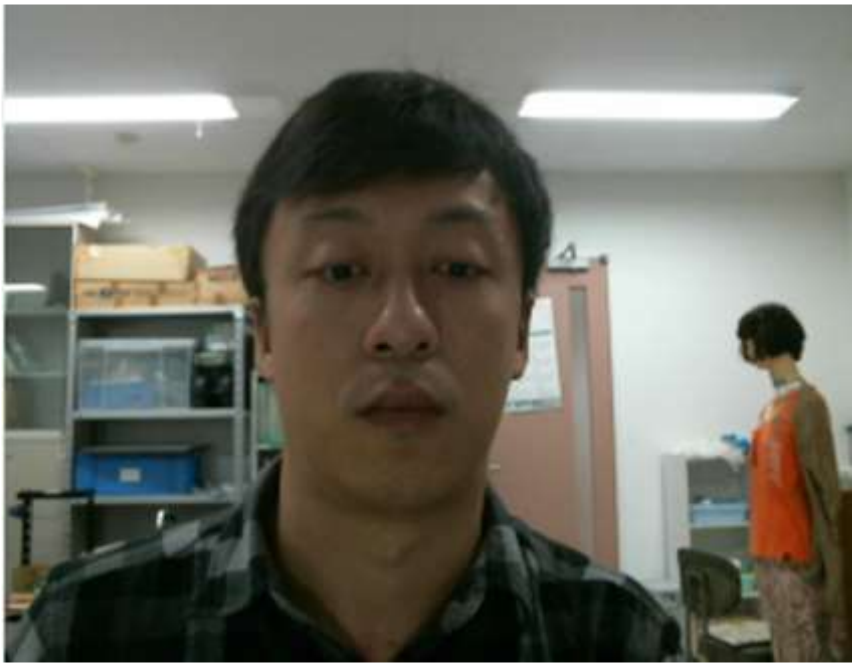
$$AR_n = \frac{x_{Rn} - x_{Ln}}{y_{Bn} - y_{Tn}} \tag{3.7}$$



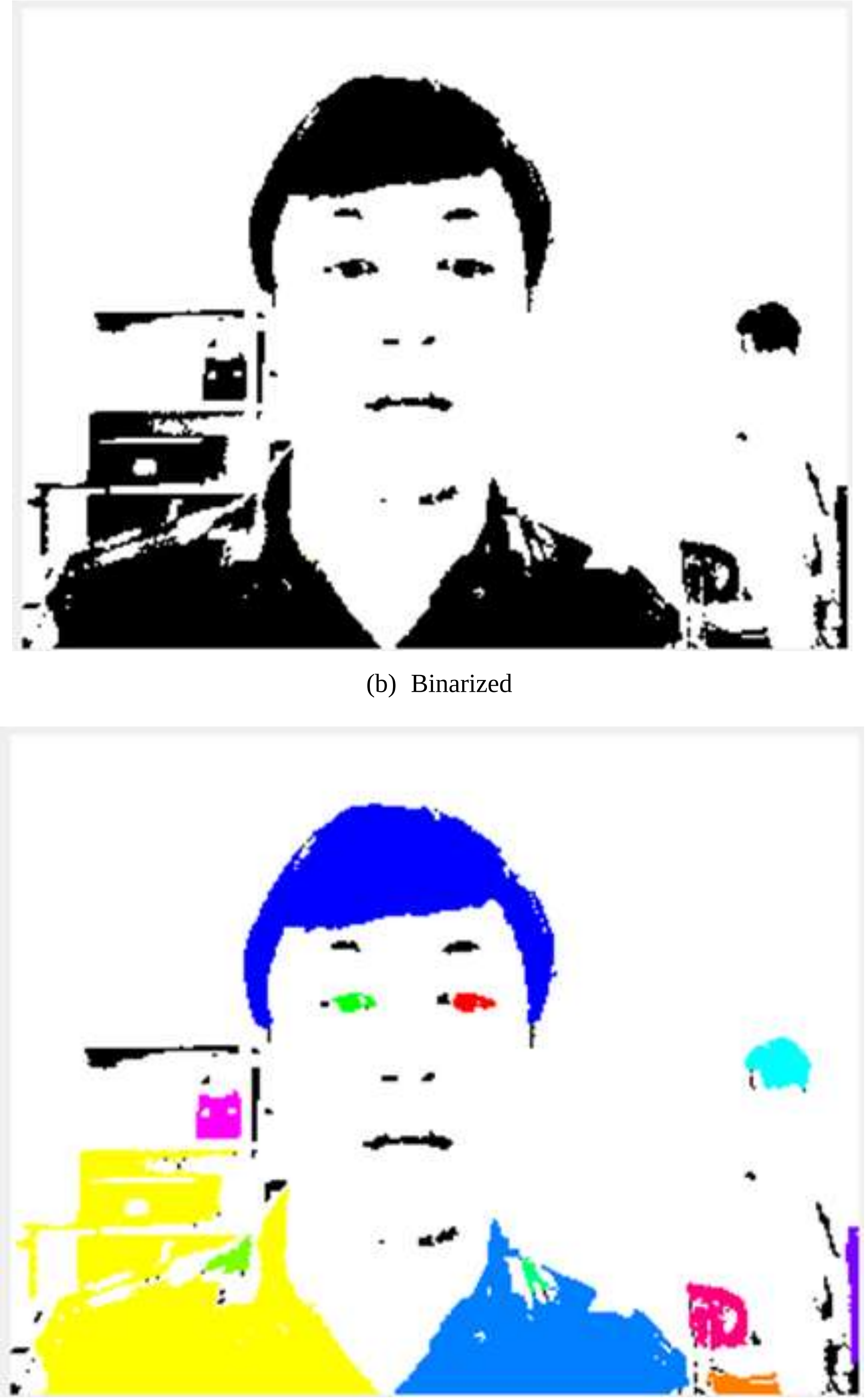
(a) Binary format image

(b) Result

Fig. 3.1 Labeling



(a) Original



(b) Binarized

(c) Labeled

Fig. 3.2 Image processing

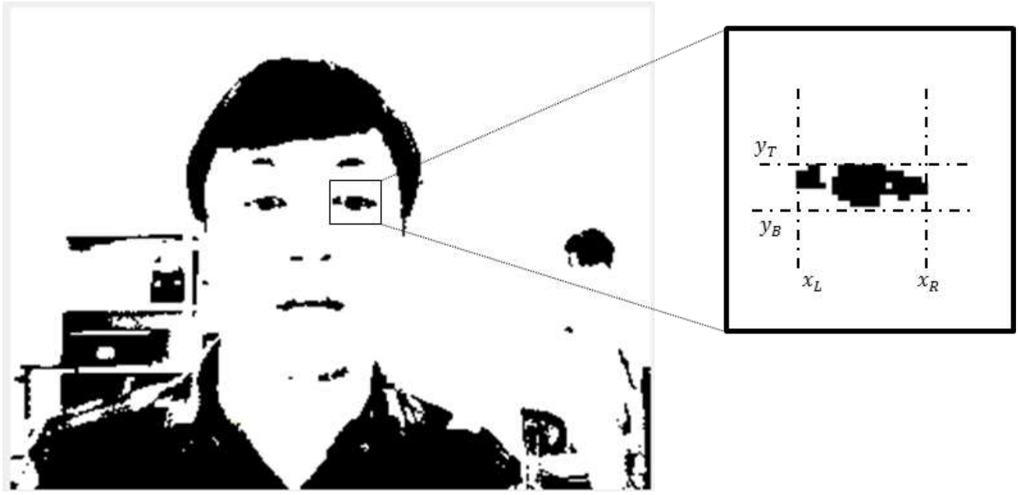


Fig. 3.3 Aspect ratio

3.2 **Basic policy of facial feature point search**

In this study, we chose nostrils as facial feature points to detect position of mouth. For the reason,

- Unlike eyes and mouth, shape is always stable.
- It is located almost in the center of the face, and there are few noises around the face.
- Since there are two, it is easy to search from features such as shape and distance.
- Since the distance between the nostrils is short, the center of the nose is stably determined and the measurement result is stabilized.

Moreover, in this system, image processing is performed by using two cameras, and since a high resolution image of 640 × 480 [pixels] is used as an input image for improving measurement accuracy, When compared with a single camera system with resolution of the input image 320 ×240 [pixels], the information processing amount per frame is eight times, and a highly efficient algorithm is indispensable for processing in real time.

In addition, the input image from the camera usually contains many unnecessary noises such as unrelated background and shadow as shown in Fig.3.2 (a). When binarizing the input image as it is as shown in Fig.3.2 (b), a large amount of noise irrelevant to the background is generated and it becomes a big hindrance to searching for the facial feature point, it is important to remove noise.

In order to solve the above problems, we decided to detect the face before detecting the nostrils. For recognition of faces, it is possible to use an image processing library such as OpenCV, but in this system we propose proprietary algorithm to realize simple and fast processing.

In this algorithm, by limiting the processing range to only the necessary part, high speed operation and noise removal are performed with low load, and it is possible to search for a definite nostril. The rough flow of the algorithm is as follows,

1. Blur the input image and extract the region close to flesh color (face) as a connected component from the blurred input image.

2. Determine the face area from the center of gravity and area of the connected component with the largest area.

3. Search nostrils within the area of the determined face.

A specific nostril search algorithm in this system will be described below.

3.3 **RGB values of input image**

In searching nostrils, we aim to realize simple and fast processing, and in this research we will search for algorithms that can automatically detect faces and nostrils in RGB color space.

First of all, it is important to detect skin. Here, RGB color information in the input image is verified as a basis for detecting the face of the user according to the color information.

For three images with the face darkness changed step by step as shown in Fig.3.4, samples of a certain number of RGB values are sampled for each part of the face and verified. The number of samples collected at each site is as shown in Table 3.1.

Fig.3.5, Fig.3.6, and Fig.3.7 are graphs plotting the color distribution in the images of Fig.3.4 (a) - (c) for each part of the face. Here, the correlation was examined for six of R, G, B, the difference between R and G, the difference between R and B, and the difference between G and B.

Also, Fig.3.8 is a plot of percentage of RGB values in each pixel on a triangular graph. Although there is a difference in brightness and darkness depending on conditions, the distribution relation of colors of RGB is almost isomorphic.

In addition, it can be confirmed that the components of *R* are stronger in the skin (◆), the nostrils (■), and the lips (▲) than in the other regions. Focusing on these three parts, almost every point holds $R>G>B$ regardless of the brightness. Among them, the difference between *R* and *B* is very large.

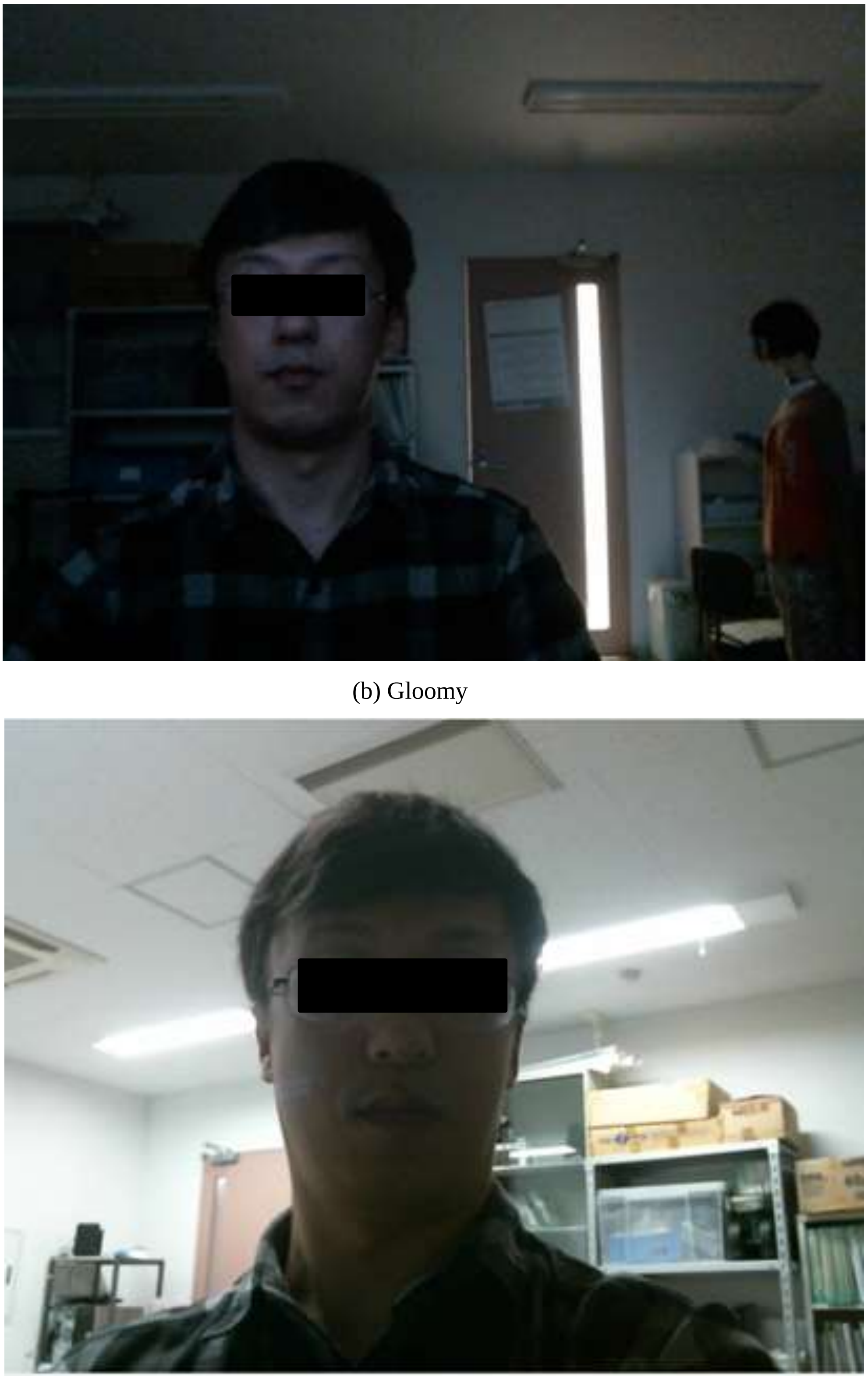
Furthermore, although it depends on the brightness, it can also be seen from Fig.3.8 that the proportion of color is almost determined by the part. Fig.3.9 only shows the skin color of Fig.3.8 (a) - (c).

Here, we paid attention to the value of *R*, *G*, *B*, the ratio between every two colors, and the occupancy of each color. By using these conditions shown in these figures, the conditions for determining the pixel as skin are determined. In this system, colors that all meet the following conditions are defined as skin color, and processing is performed as a face candidate.

$$\begin{cases} R<160 \\ G<120 \\ B<90 \\ R\text{-}G>0 \\ R\text{-}1.5B>0 \\ G\text{-}B>0 \\ 0.38<\dfrac{R}{R+G+B}<0.6 \\ 0.27<\dfrac{G}{R+G+B}<0.49 \\ 0.02<\dfrac{B}{R+G+B}<0.3 \end{cases} \tag{3.8}$$

[redacted]

(a) Well-lighted



[redacted]

(b) Gloomy

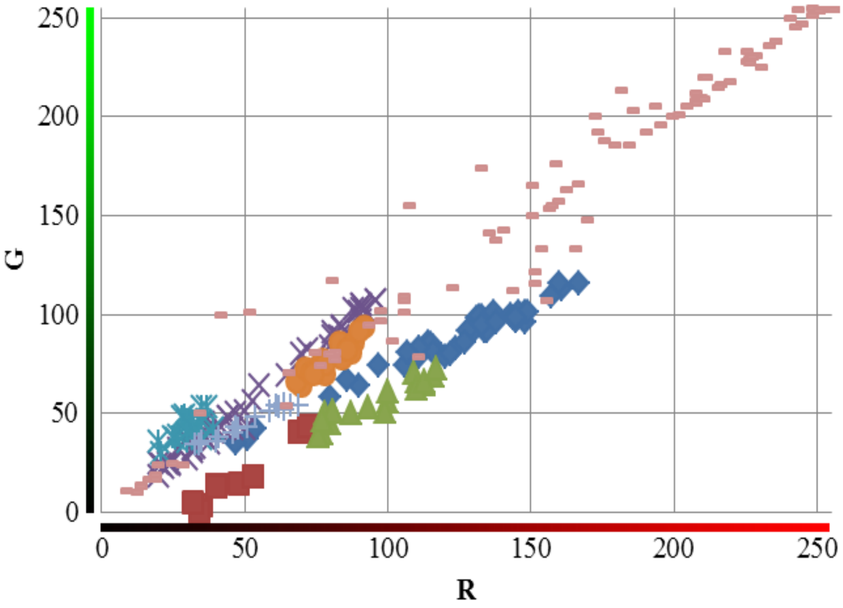
[redacted]

(c) Backlight

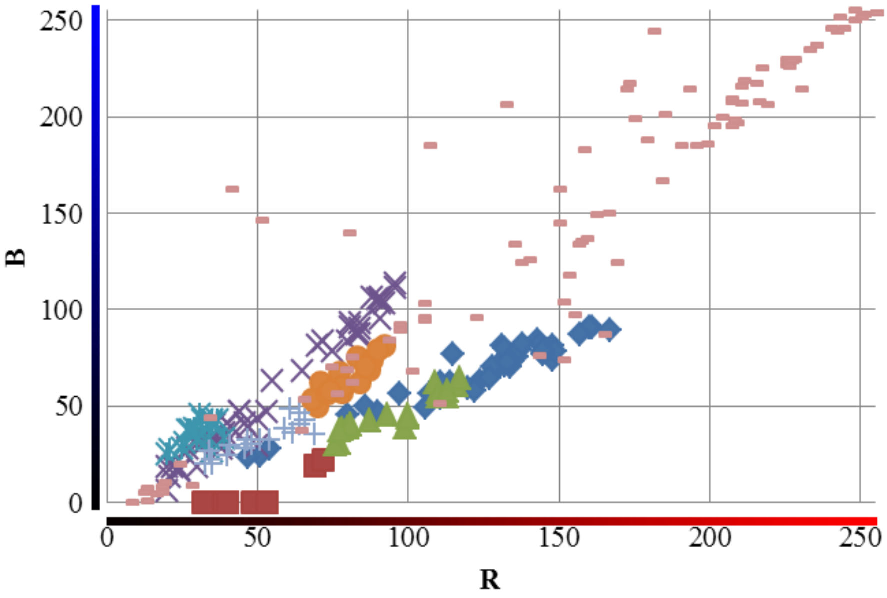
Fig.3.4 Color sample image

Table.3.1 Number of sample

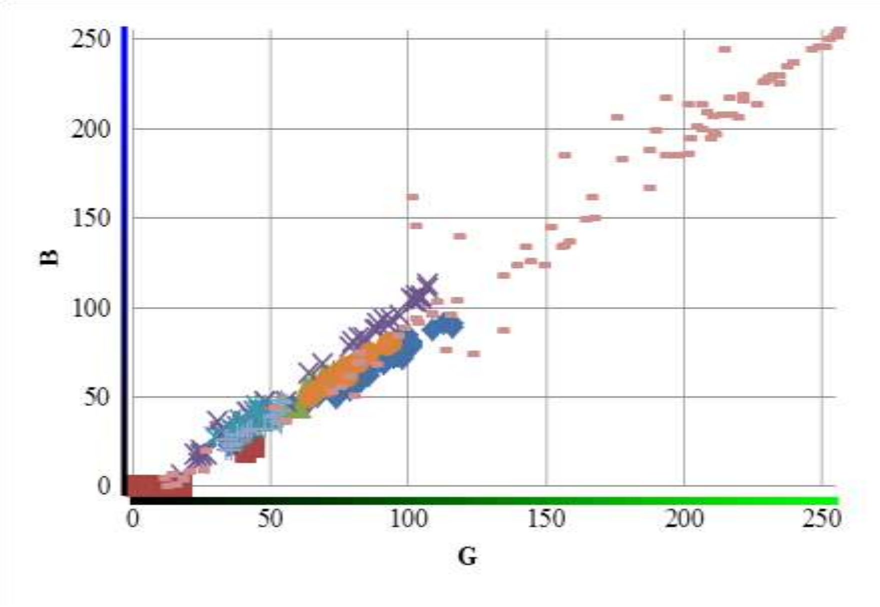
Part	Number of sample
Skin	50
Nostril	20
Lip	20
Hair	50
Pupil	10
White of the eye	10
Eyebrow	20
Background	100


G
250
200
150
100
50
0
0
50
100
150
200
250
R

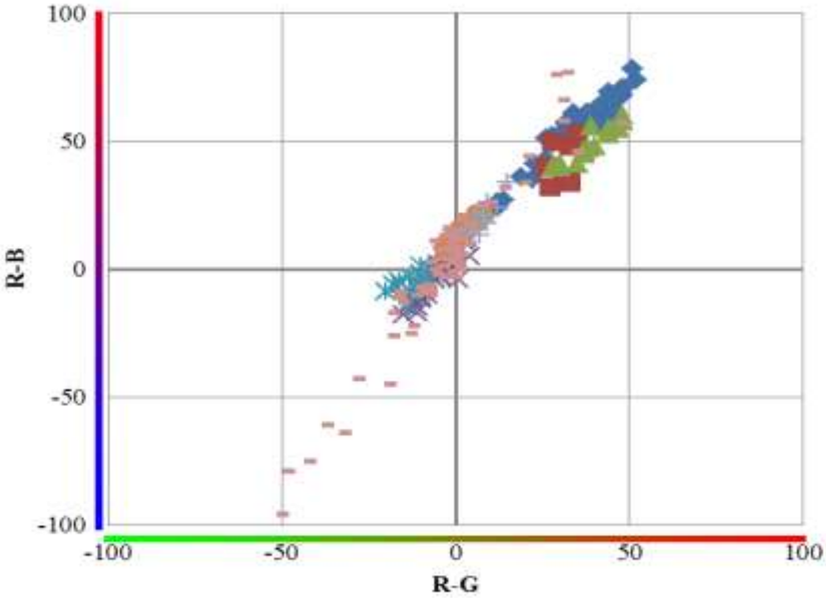
(a) R-G


B
250
200
150
100
50
0
0
50
100
150
200
250
R

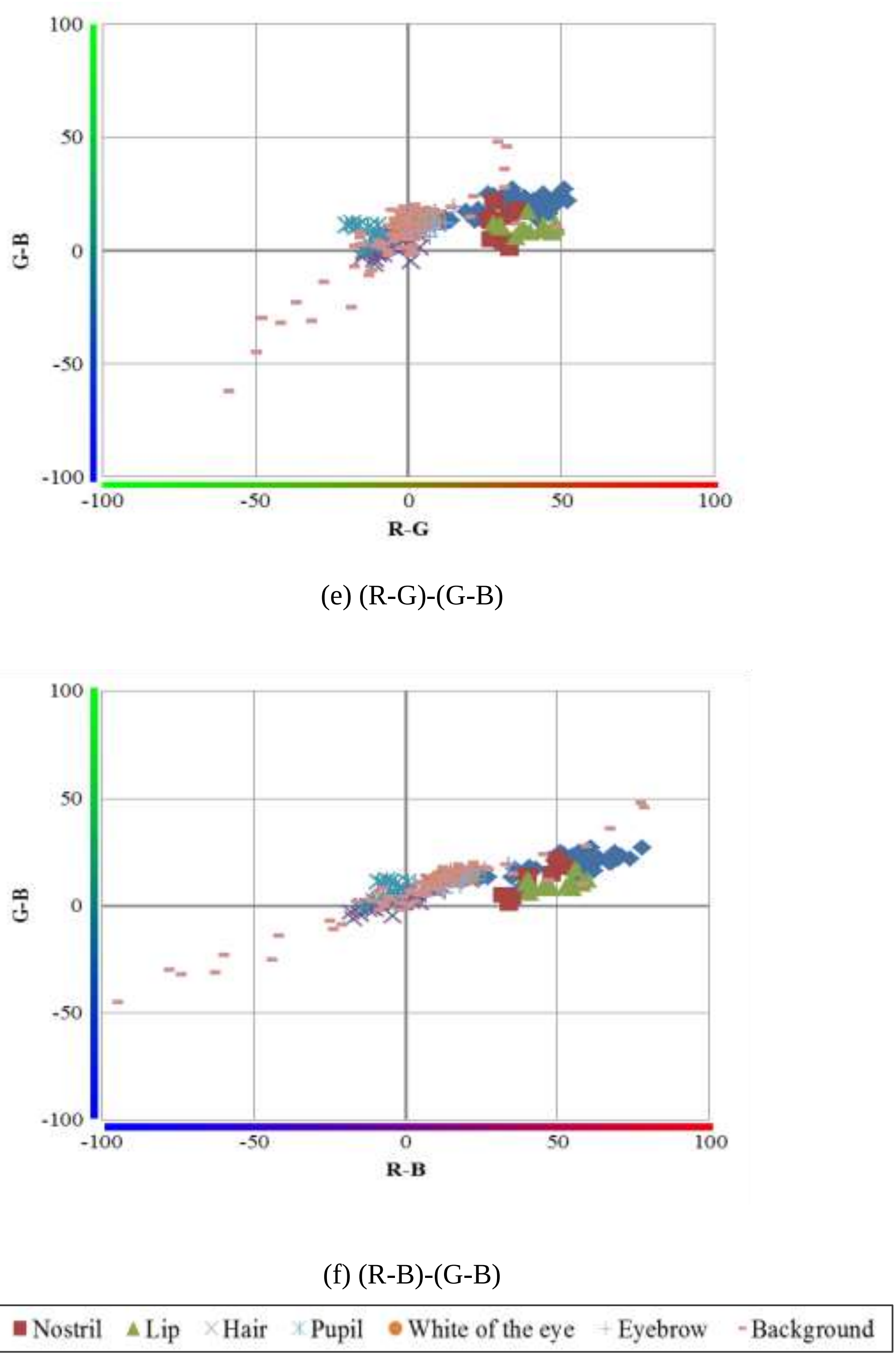
(b) R-B


250
200
150
100
50
0
B
0
50
100
150
200
250
G

(c) G-B


100
50
0
-50
-100
R-B
-100
-50
0
50
100
R-G

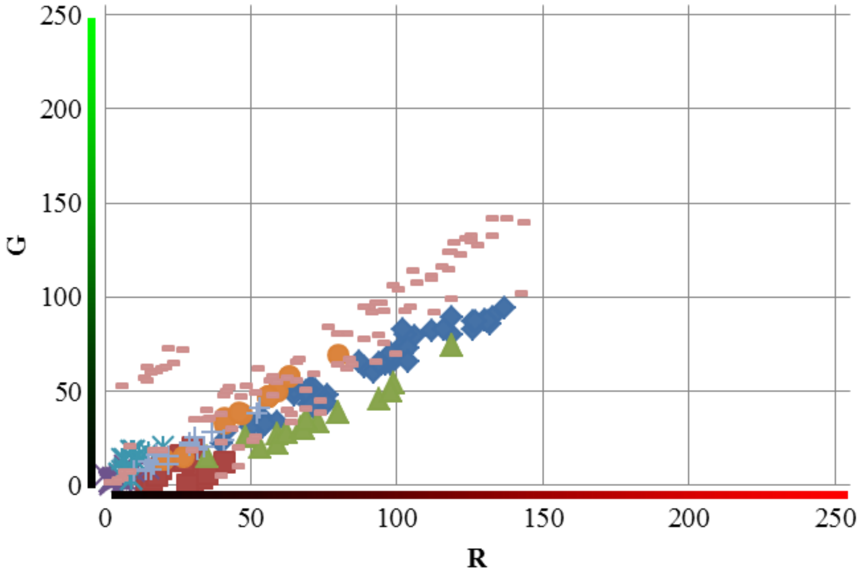
(d) (R-G)-(R-B)



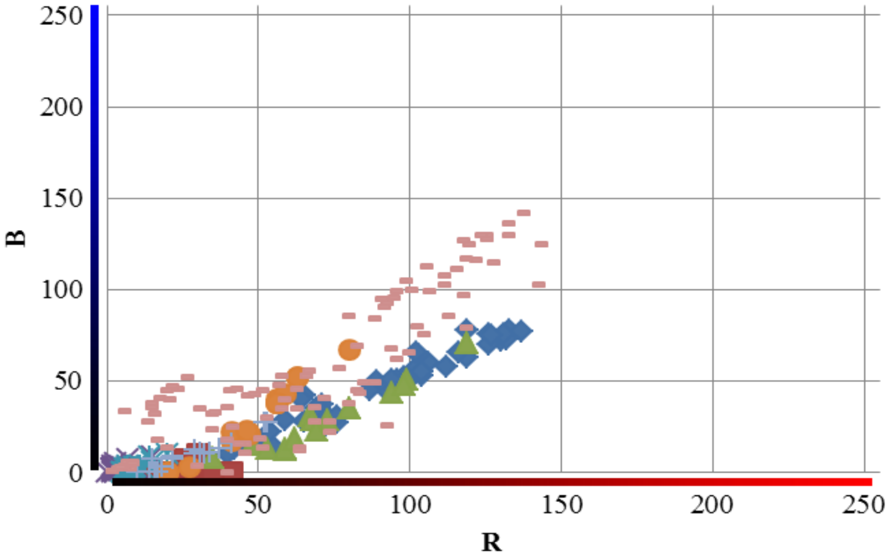
(e) (R-G)-(G-B)

(f) (R-B)-(G-B)

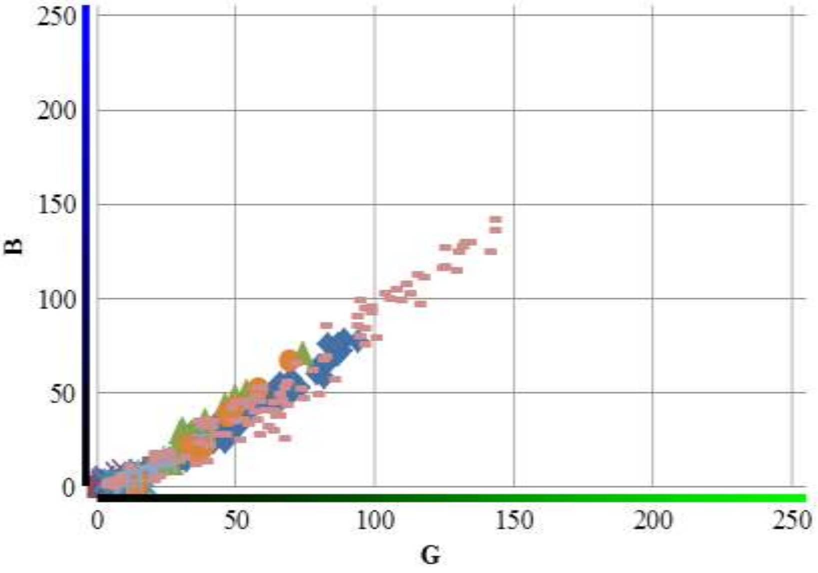
Fig. 3.5 Color of the parts of a face (Well-lighted)


G
250
200
150
100
50
0
0
50
100
150
200
250
R

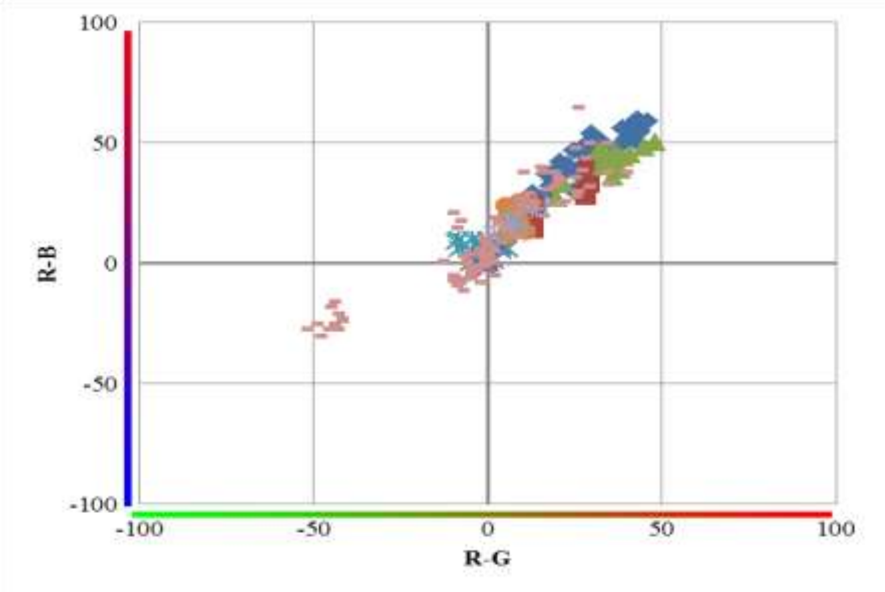
(a) R-G


B
250
200
150
100
50
0
0
50
100
150
200
250
R

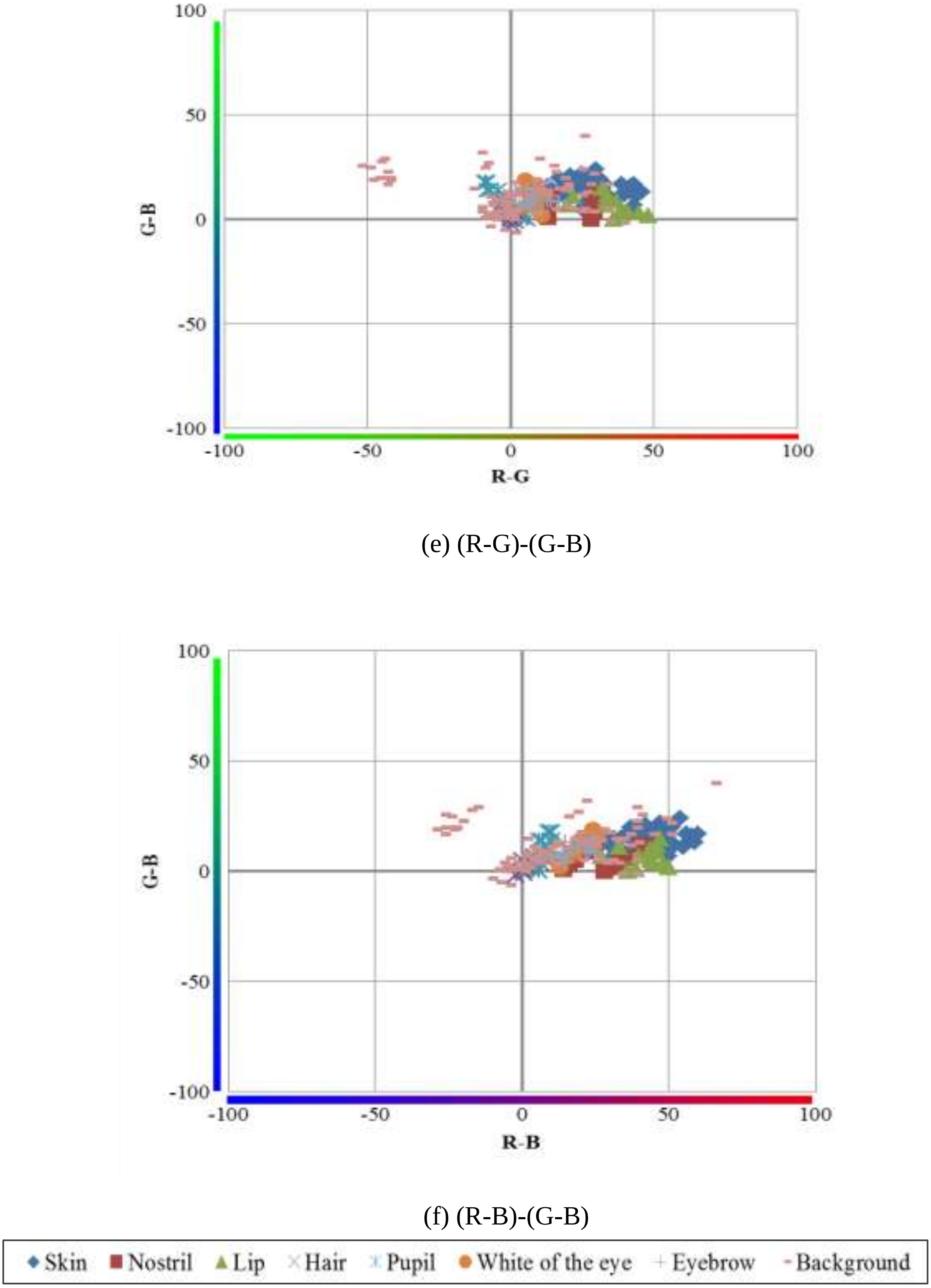
(b) R-B


250
200
150
100
50
0
B
0
50
100
150
200
250
G

(c) G-B


100
50
0
-50
-100
R-B
-100
-50
0
50
100
R-G

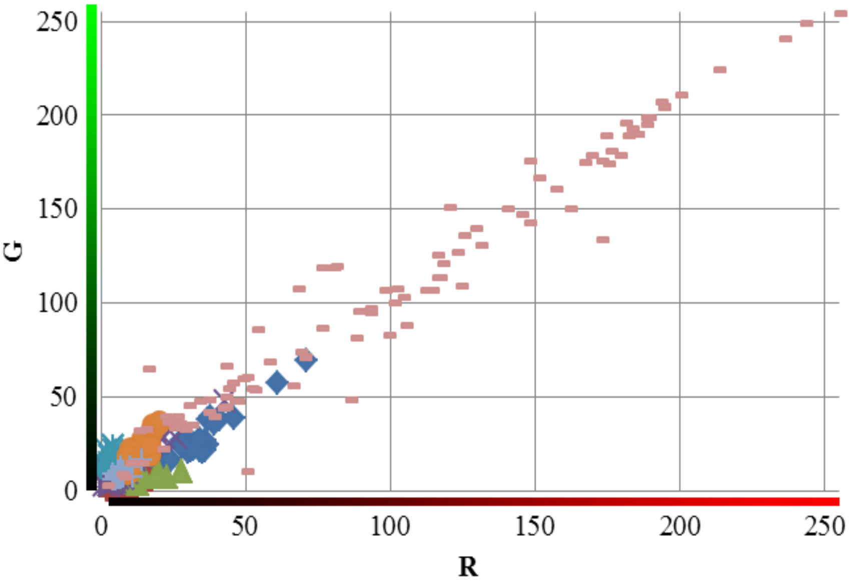
(d) (R-G)-(R-B)



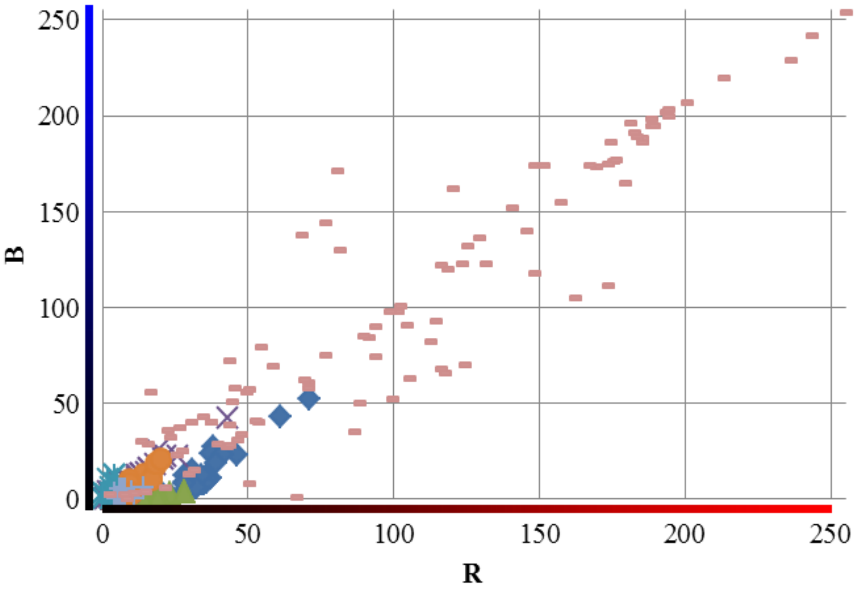
(e) (R-G)-(G-B)

(f) (R-B)-(G-B)

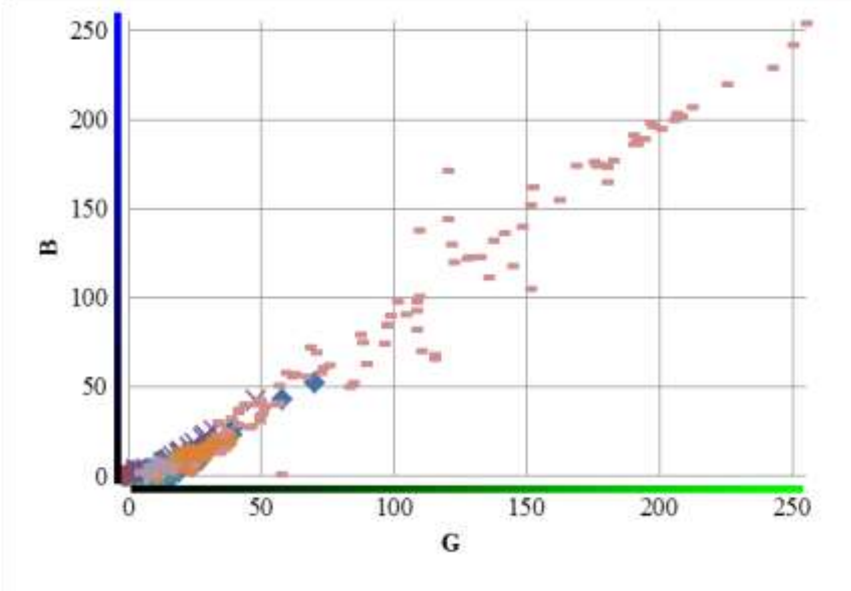
Fig. 3.6 Color of the parts of a face (Gloomy)


250
200
150
100
50
0
G
0
50
100
150
200
250
R

(a) R-G

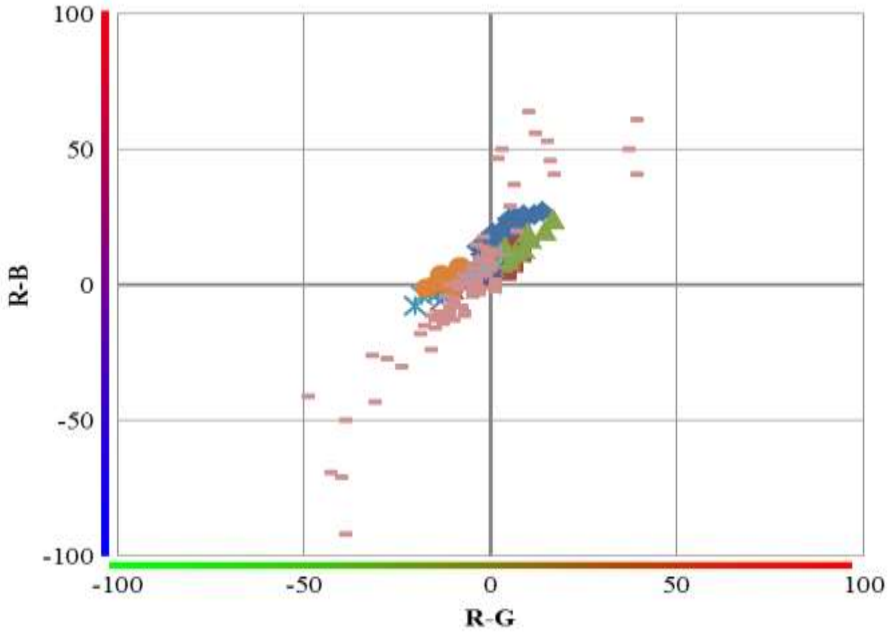

250
200
150
100
50
0
B
0
50
100
150
200
250
R

(b) R-B



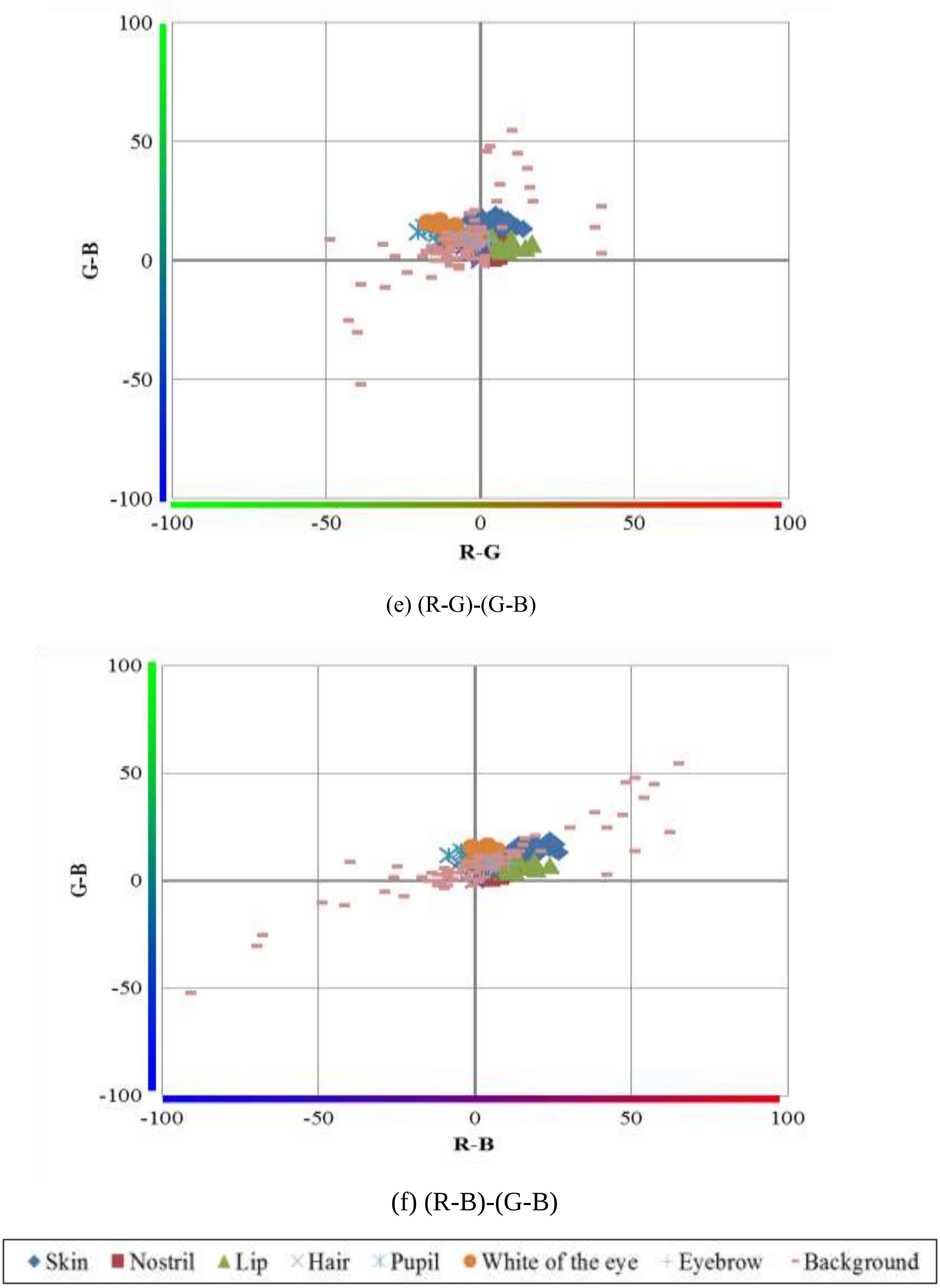
B
250
200
150
100
50
0
0
50
100
150
200
250
G

(c) G-B



100
50
0
-50
-100
R-B
-100
-50
0
50
100
R-G

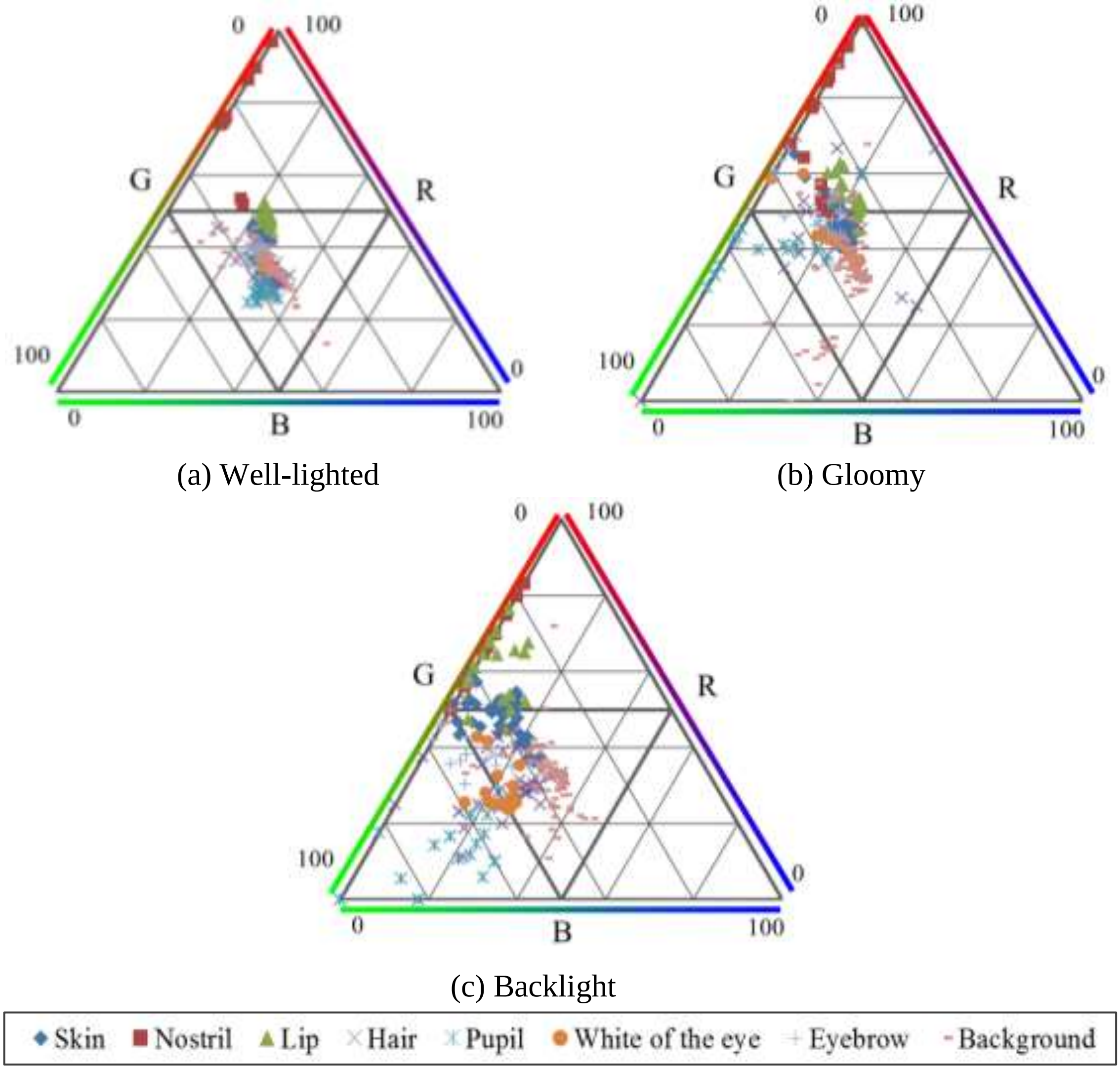
(d) (R-G)-(R-B)



(e) (R-G)-(G-B)

(f) (R-B)-(G-B)

Fig. 3.7 Color of the parts of a face (Backlight)



(a) Well-lighted

(b) Gloomy

(c) Backlight

Fig. 3.8 Percentage of RGB

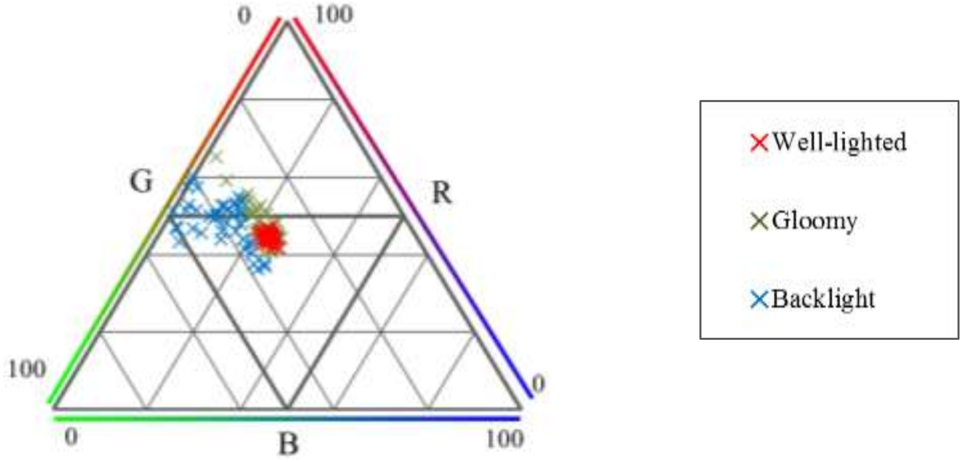


Fig. 3.9 Percentage of RGB (Skin)

3.4 **Face detection algorithm**

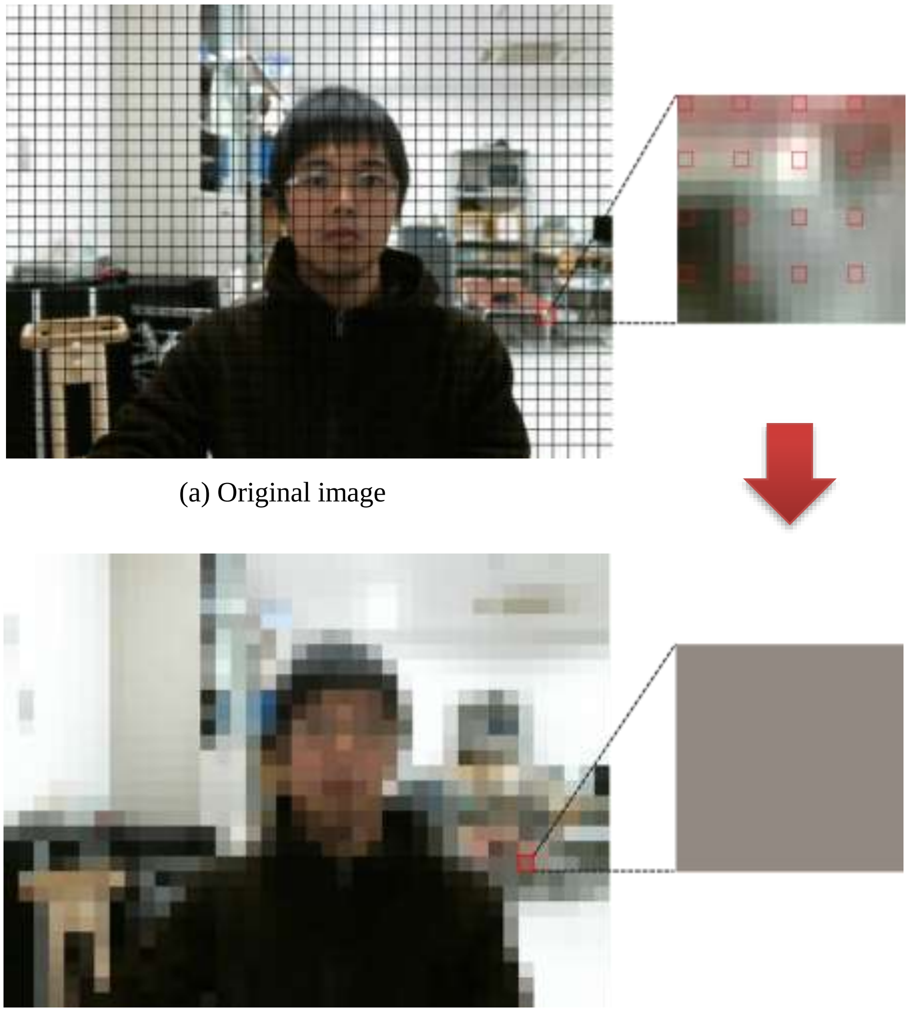
When detecting a face by processing, it is important to recognize the entire face as one connected component, but it is often not recognizable as a connected component depending on the state of the face such as shadows and spectacles. Therefore, in this algorithm, processing is performed on a blurred image of the original image in a mosaic pattern.

With 16 × 16 [pixels] as one block, divide 640 × 480 [pixels] into 40 × 30 blocks. Each block representative color (R',G',B') was determined from the average color of the representative points as follows.

$$(R',G',B')=\left(\sum_{i=0}^{3}\sum_{j=0}^{3}\frac{R[4i,4j]}{16},\sum_{i=0}^{3}\sum_{j=0}^{3}\frac{G[4i,4j]}{16},\sum_{i=0}^{3}\sum_{j=0}^{3}\frac{B[4i,4j]}{16}\right) \qquad (3.9)$$

This size is based on the occupancy of human face with input image, here we determined to use at least ten blocks to express human face in horizontal direction. So we devided input image by 16×16 [pixels].

The result is shown in Fig3.10. Where, $R[i,j], G[i,j], B[i,j]$ are the luminance of *R*, *G*, *B* at coordinates (i,j).



(a) Original image

(b) Result

Fig. 3.10 Pixelization

For the image in Fig.3.10 (b), binarization is performed on the region satisfying the expression (3.8) and extracted, which is shown in Fig.3.11. In this state, it is easier to be affected by noise due to light and dark, so we decided to smooth the evaluation. In this case, as shown in Fig.3.12, a method of adding to eight adjoining blocks is adopted.

Let sgn $g[i, j]$ in the coordinates (i, j) be the number of points after smoothing $g'[i, j]$,

$$g'[i, j] = \sum_{k=i-1}^{i+1} \sum_{k=j-1}^{j+1} g[i, j] \tag{3.10}$$

The result of smoothing processing on the image in Fig.3.11 is shown in Fig.3.13. In Fig. 3.13 (a), the evaluation is higher for areas with redness, and Fig.3.13 (b) shows the area re-extracted the area where the evaluation is above a certain level. In this example, the extraction of $g'[i, j] \geq 1$ is performed, but by narrowing this threshold it is also possible to narrow the range. Subsequently, components having the largest area among these extracted connected components are selected, and processing is performed for that area. This is shown in Fig. 3.13 (c).

Next, this area is converted into a rectangle with length and width of 4: 3. As shown in Fig. 3.14, if the area S of the area is taken as the center of gravity(G_x, G_y), $k = \sqrt{S/12}$ let's be the rectangle of height $4k$ and width $3k$.

Further, binarization processing is performed on the interior of this rectangle using equation (3.8), skin is extracted again, and the area and the center of gravity are determined to determine the area of where the face was finally.

The basic flow is the same as the first extraction, so it is omitted, but for the second time we have done processing on all pixels and not labeling processing. The results of binarization and the final results are shown in Fig. 3.15 and Fig.3.16. The red frame in the figure is the area obtained by the first search, and the white frame is the final area. From the fact that the center of gravity of the rectangle is closer to the center of the face than the first search and the background noise contained in the left side is also cut, it can be understood that the area of the face can be further narrowed down.

Fig.3.17 compares the binary image of the original image and binarized image after determining face area by this algorithm. Fig.3.17 (b) shows that the noise has been removed considerably, and the effectiveness of the algorithm proposed in this research can be confirmed.

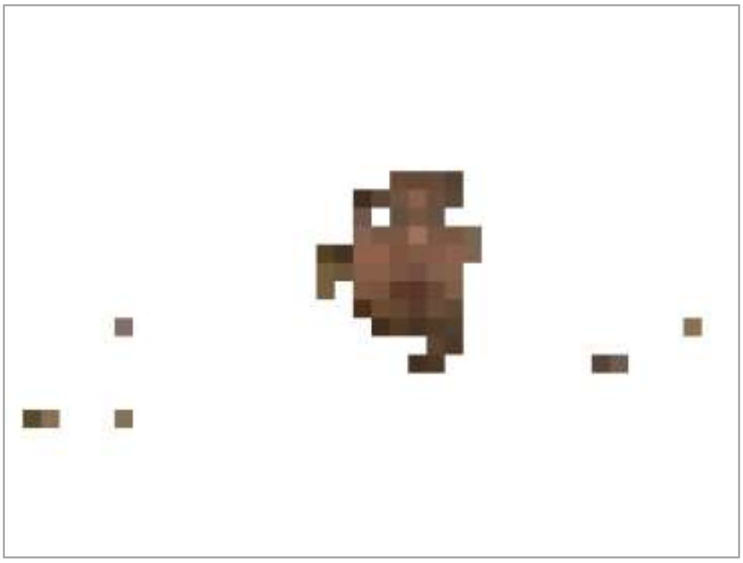


Fig. 3.11 Binarization result

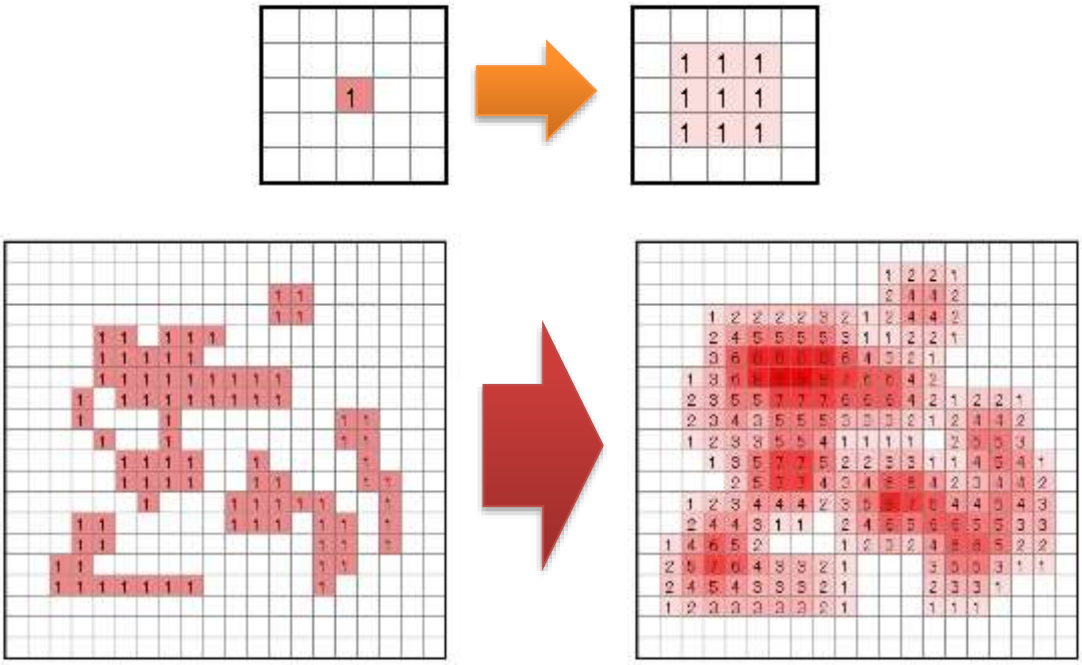
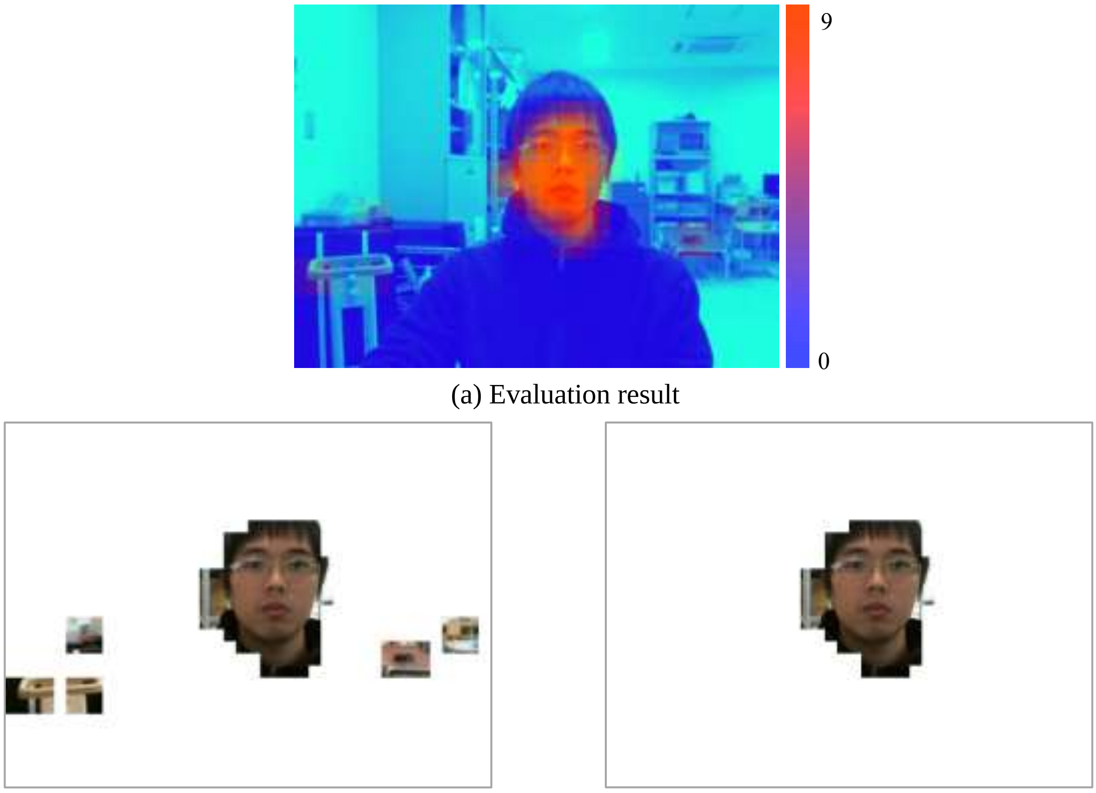


Fig. 3.12 Smoothing

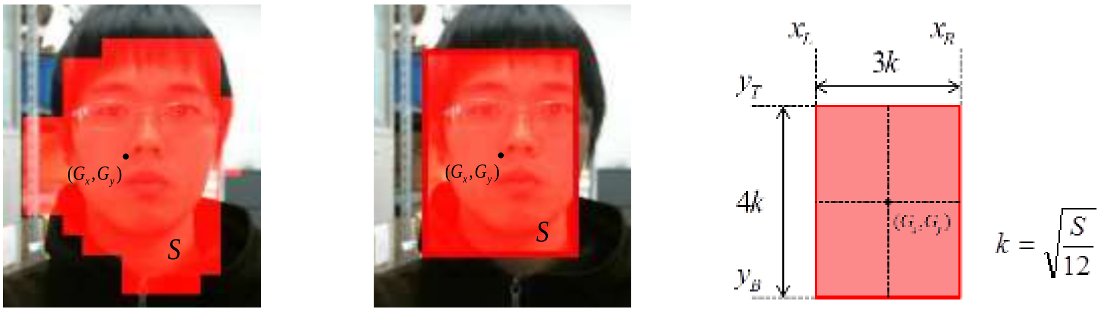


(a) Evaluation result

(b) Region extraction

(c) Choose biggest one

Fig. 3.13 Facial recognition



(a) Original region

(b) Rectangular region

(c) Side length

Fig. 3.14 Zone assignment

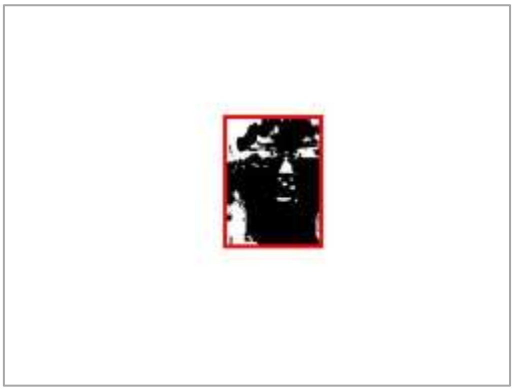


Fig. 3.15 Binarization result

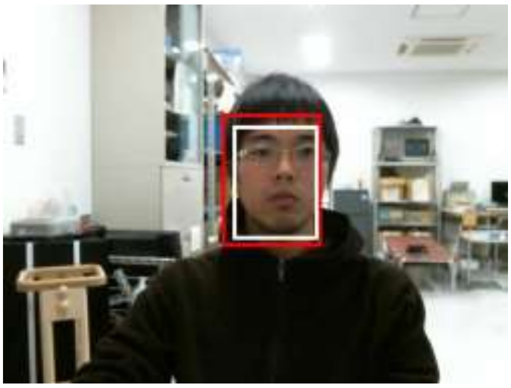
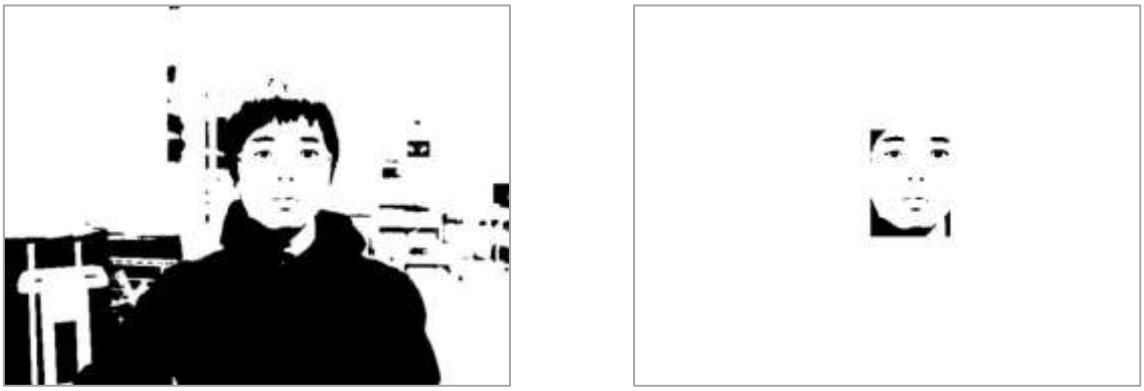


Fig. 3.16 Final result



(a) Original image

(b) With facial recognition

Fig. 3.17 Noise reduction

3.5 Verification experiment

3.5.1 Comparative experiment between our system and OpenCV

In order to verify the effectiveness of this algorithm, face detection is carried out using an actual image as input data. Eight images of 640 × 480 [pixels] still images whose conditions such as ambient brightness and face angle are changed are used as input images, and faces are detected by their respective algorithms.

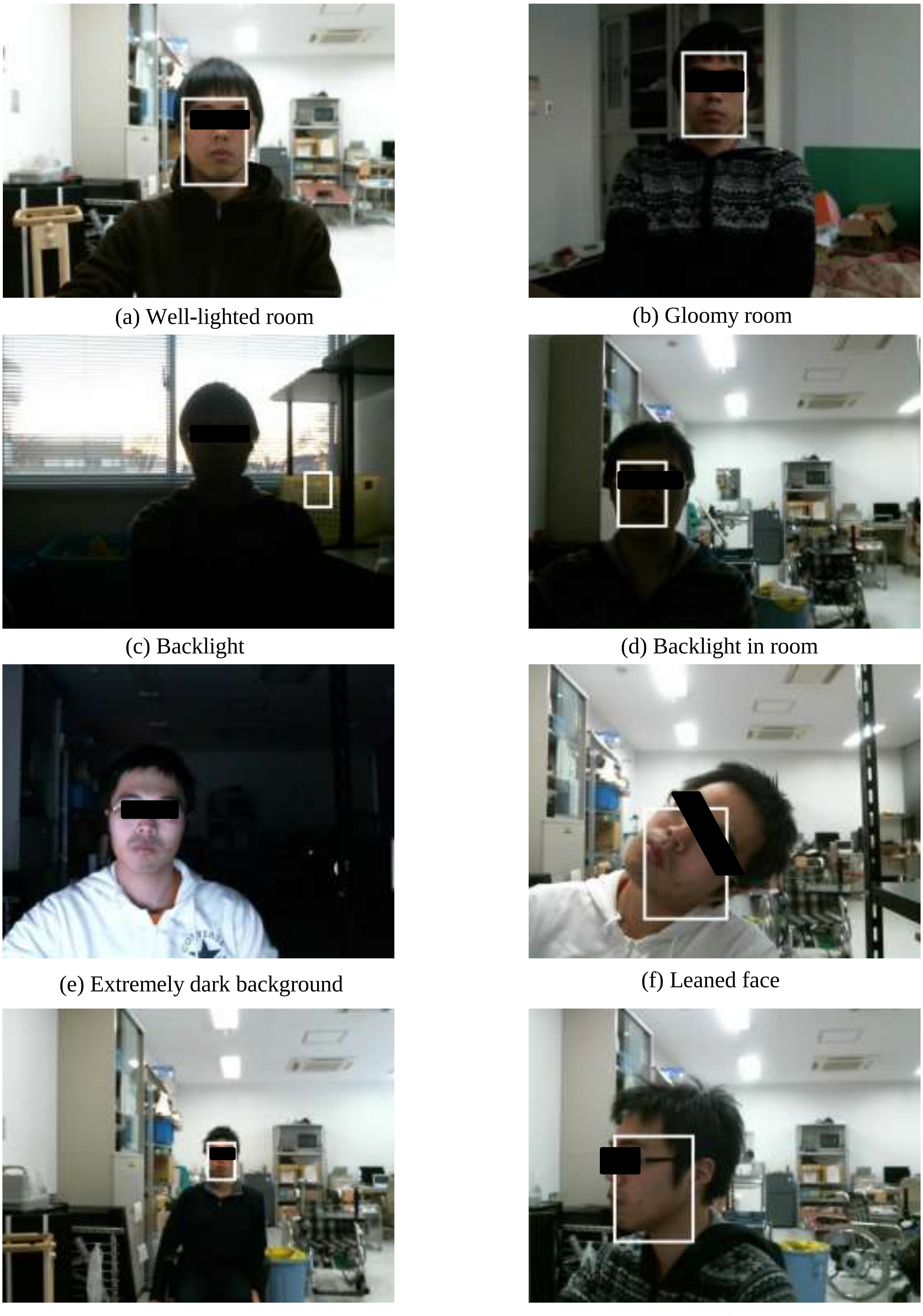
As a comparison with this algorithm, we also decided to perform similar processing with OpenCV face detection function. OpenCV is a library for open source computer vision developed and published by Intel and is widely used in the field of image processing. Face detection by OpenCV is based on Viola-Jones method [13].

The result of face detection using the algorithm described in this research is shown in Fig. 3.18, and the face detection using the OpenCV library is shown in Fig.3.19.

In the case of propose method, in the condition such as well-lighted, gloomy, and backlight, our system can detect human face successfully. Even though the face is tilted, and to the side, the system can still detect human face. If the human is in a long distance from the camera, the system can detect human face. However, in the condition such as backlight and extremely dark background, our system can not detect human face.

The recoginition by OpenCV failed as shown in Fig.3.19 (d), (f), and (h), but the detection of our research was successful as shown in Fig.3.18 (d), (f), and (h), so it can be confirmed that it is very effective in this system on the premise of face detection.

So when the light is dark, sometimes, our system can not detect human face, but our system is developed for meal support equipment, so the condition of light is always bright , and when the light is bright enough, our system is better than OpenCV in face detection. It proves that our system is useful for meal support equipment.



(a) Well-lighted room

(b) Gloomy room

(c) Backlight

(d) Backlight in room

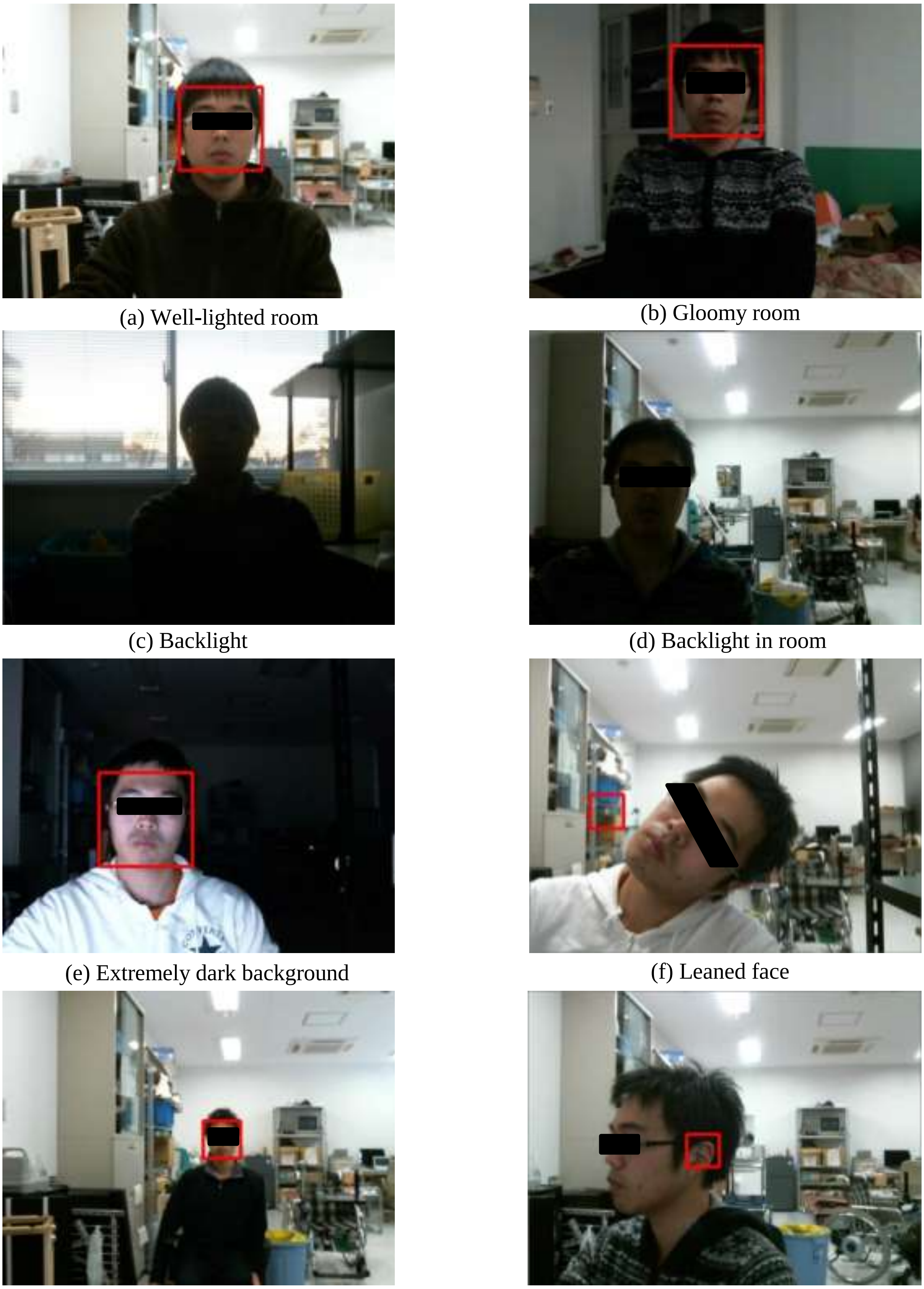
(e) Extremely dark background

(f) Leaned face

(g) Distant position

(h) Profile

Fig. 3.18 Face Recognition (This algorithm)



(a) Well-lighted room

(b) Gloomy room

(c) Backlight

(d) Backlight in room

(e) Extremely dark background

(f) Leaned face

(g) Distant position

(h) Profile

Fig. 3.19 Face Recognition (OpenCV)

3.5.2 Comparative experiment between our system and other camera equipment

Because of there are kinds of face detection in our daily life, we compare face detection system with the other detection system, such as camera (FINEPIX Z-200fd of FUJIFILM corporation), and smartphone (IPhone X of Apple Inc.). The Viola-Jones method is generally applied to detection of human face by the camera and smartphone. The result is shown in Table.3.2.

In this table, the circle (○) means that the face of subject could be detected successfully, the triangle (△) means that the system mistake the face for the other part of the image, and the crosses (×) means that nothing was detected in input image.

We can see that accuracy rate of our system is almost 100%, and in front side, the accuracy rate of camera and smartphone is also 100%, however, in left and right side by camera, the accuracy rate is 33.3% and mismeasurement rate is 25%, in tile side by camera, the accuracy rate is 41.6%, and in left and right side by smartphone, the accuracy rate is 29.1%, in tile side by smartphone, the accuracy rate is 58.3%, so in face detection, our system is better than the other method.

Table.3.2 comparative experiment

	Our system				Camera (FINEPIX)				Smartphone (IPhone X)			
Subject	front	left	right	tilt	front	left	right	tilt	front	left	right	tilt
A	○	○	○	○	○	×	×	×	○	×	×	○
B	○	○	○	○	○	×	○	×	○	×	×	○
C	○	○	○	○	○	×	○	○	○	○	×	○
D	○	○	○	○	○	△	×	×	○	×	×	○
E	○	○	○	○	○	×	△	○	○	×	○	○
F	○	○	○	○	○	○	○	○	○	×	×	×
G	○	○	○	○	○	×	○	×	○	×	×	×
H	○	○	○	○	○	△	×	×	○	×	○	×
I	○	○	○	○	○	○	○	×	○	○	○	○
J	○	○	○	○	○	×	△	○	○	×	×	×
K	○	○	○	○	○	△	×	×	○	×	×	×
L	○	○	○	○	○	○	△	○	○	○	○	○

○ Detected △ Misrecognized ×Not detected

3.5.3 Setting the evaluation function

Focusing on the fact that there are many dark areas around the eyes, mouth, and chin, while paying attention to the fact that the proportion of bright skin around the nostrils is very large, the proportion of the skin occupied by each value y in the face area is investigated, here the value y means every row in transverse direction and as an evaluation function I decided to use it. When the width of the face area is x [pixel], the ratio occupied by the skin in y.

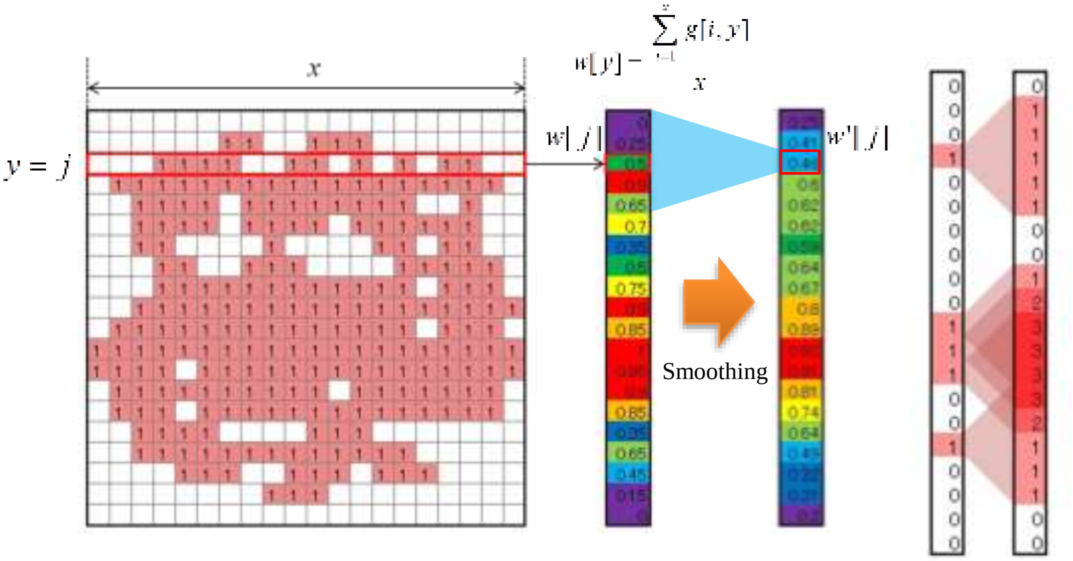
$$w[y] = \frac{\sum_{i=1}^{x} g[i, y]}{x} \tag{3.11}$$

This is smoothed and the noise removed is used for determining the nostrils. Assuming that the influence range in smoothing is up to the upper and lower n rows,

$$w'[y] = \frac{\sum_{j=y-n}^{y+n} w[j]}{n} \tag{3.12}$$

In this system, n is 80, it means that the area of face was divided into 80 rows. The flow of this series of processing is shown in Fig.3.20, and the result of actually processing the input image is shown in Fig.3.21.

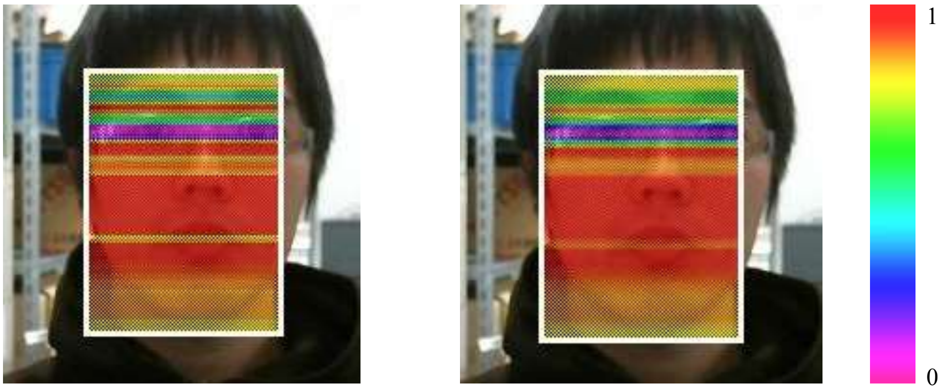
While the eyes and eyebrows are purple and blue, the mouth and chin are nearly yellow in color, the nostrils are red and the percentage of skin is very high. So we can reduce the area of nostril detection and accelerate the detection speed.



(a) Weighting process

(b) Smoothing

Fig. 3.20 Weighting



(a) Weighting

(b) Smoothing

Fig. 3.21 Weighting result

3.6 Nostril detection algorithm

3.6.1 Removal of unnecessary connected components

Detection of a nostril in the face region obtained by the method in the previous section will be described below.

Fig.3.22 (a) shows the result of binarizing and labeling the face region, but it can be confirmed that many connected components remain in addition to the nostrils. Therefore, narrowing down is performed from the feature quantity of each connected component.

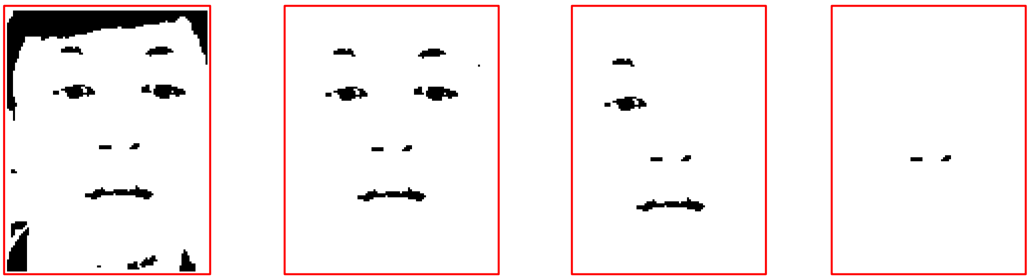
First of all, it is a connected component that has a point on a side of a rectangle representing a face area.

The connected component existing at the end of the region is only one end of the component originally connected to the outside of the region, and its shape and size become distorted, which is likely to cause erroneous detection. Also, we excluded from the fact that the nostrils to be searched existed near the center of the face and judged that there was no problem. As a result, it becomes as shown in Fig.3.22 (b).

Next, narrow down according to the size of the component. Verification in the range of 300 to 1000 mm in this system confirmed that the area per nostril fell within the range of approximately 10 to 200 [pixels]. Therefore, connected components having areas not included in this range are excluded. It can be confirmed in Fig. 3.22 (c) that very small noise and components of large size are excluded.

Further, narrowing down is performed by the aspect ratio of expression (3.7). We decided to exclude components that do not satisfy the conditions $1/3 < AR < 3$ in this system. Horizontal mouth and eyes etc. are removed in Fig.3.22 (d).

The connected component left by narrowing down above is set as a nostril candidate. In the example of Fig.3.20, only the nostril remains, but of course there are often more than three components left. Therefore, when there are n remaining components, it is necessary to verify all the combinations ${}_nC_2$ and narrow the components to one pair using the evaluation function.



(a) Binarization (b) Border (c) Size (d) Aspect

Fig. 3.22 Noise elimination method

3.6.2 Filtering nostril candidates

Let the area S_1 , S_2 of the two components be the center of gravity $G_1 = (x_1, y_1)$, $G_2 = (x_2, y_2)$ an example is shown in Fig.3.23.

First of all, the area ratio is used as a feature quantity for detecting pairs of nostrils. Since the area of the nostrils is not greatly different between the left and right, only the pairs of connected components which satisfy the $1/3 < S_1/S_2 < 3$ are verified, and the rest are excluded.

Next is the slope between the nostrils. The inclination of the straight line passing through the two points is,

$$\frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} \tag{3.13}$$

We narrowed down by the range of this slope. In this research it was made $-0.5 < \Delta y / \Delta x < 0.5$.

For those that satisfy the above two conditions, then the distance between each connected component is confirmed. In this research we will use the distance between centroids.

If we let the centroid of the two components be $G_1 = (x_1, y_1)$, $G_2 = (x_2, y_2)$ then that distance d is

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{3.14}$$

. In the camera of this system, it was confirmed $d < 50$ by verification that it is at a position of about 300 [mm] or more, so we will use this condition also for narrowing down.

At this stage,

- The size of the component is close
- The inclination of the straight line passing through two points is small
- Short distance between two points

Therefore, the evaluation function with the connected component m and n is

$$e[m,n] = (10\left|\frac{\Delta y}{\Delta x}\right| + d) \times (1 - w[y_m]) \times (1 - w[y_n]) \tag{3.15}$$

We decided to recognize $e_{\min} = e[m,n]$ as m,n as a nostril. By this calculation, it is possible to withdraw the nostril preferentially from the connected component with high evaluation. However, the center of gravity of m and n is $G_m = (x_m, y_m)$, $G_n = (x_n, y_n)$.

The final result is shown in Fig.3.24. With this method it is possible to reliably detect both nostrils, we can calculate the midpoint $(x, y) = (\{x_1 + x_2\}/2, \{y_1 + y_2\}/2)$ and use it as input to the position measurement system, it is possible to measure the position of the face in three-dimensional space with high speed and high accuracy.

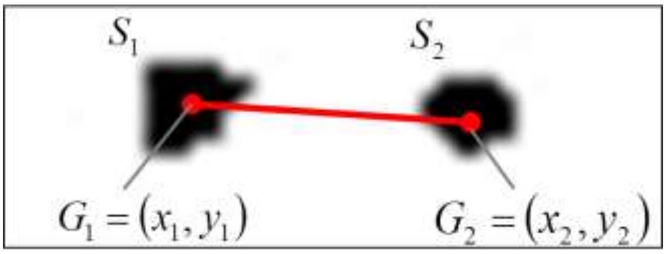


Fig. 3.23 Feature value

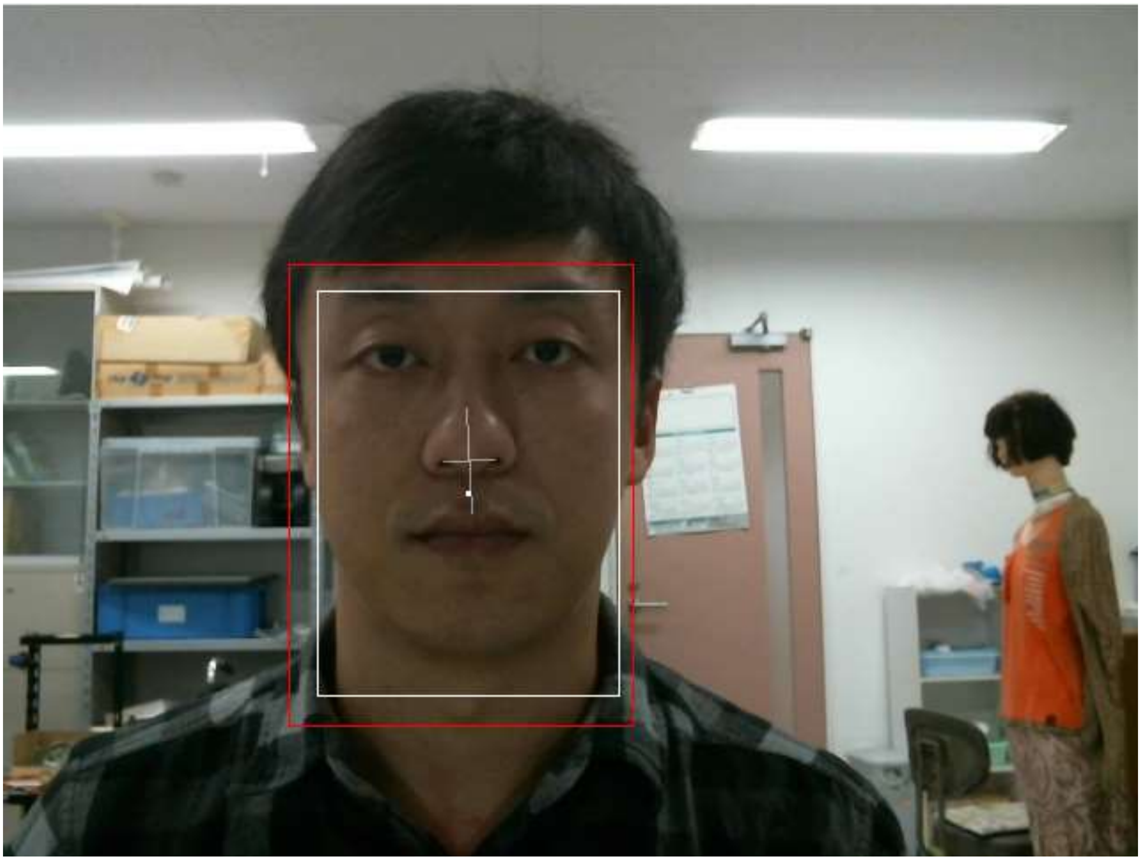


Fig. 3.24 Final result

3.7 **The relationship between mouth and nostril**

In this research, the purpose is to carry food into human mouth, so after we detected the position of nostril because the shape of the nostril is stable compare with the other parts of the face, we pay attention to the position of the mouth according to the relative position of the mouth with respect to the nostril. Fig.3.25 shows the result of binaryzation about mouth and nostril, and as shown in Fig3.26, the position of mouth always under the position of nostril, so in this research, we scan the place after the binaryzation with biggest area under the nostrils to locate the position of the mouth.

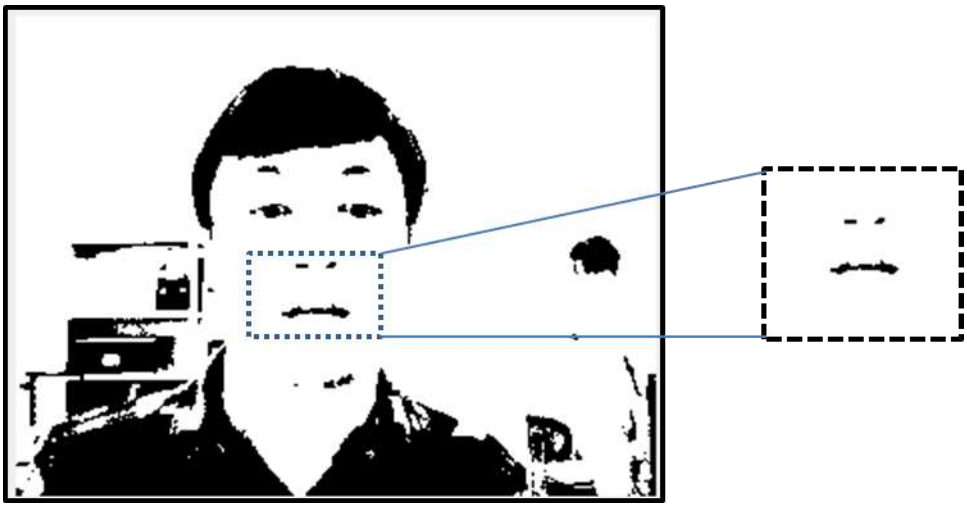


Fig.3.25 Binaryzation of mouth and nostril

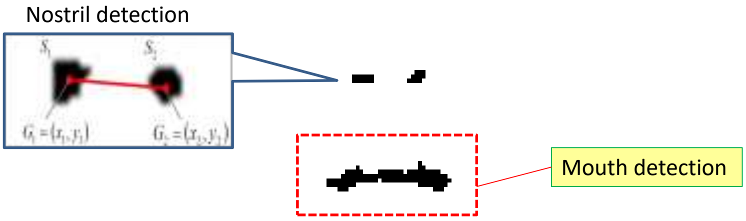
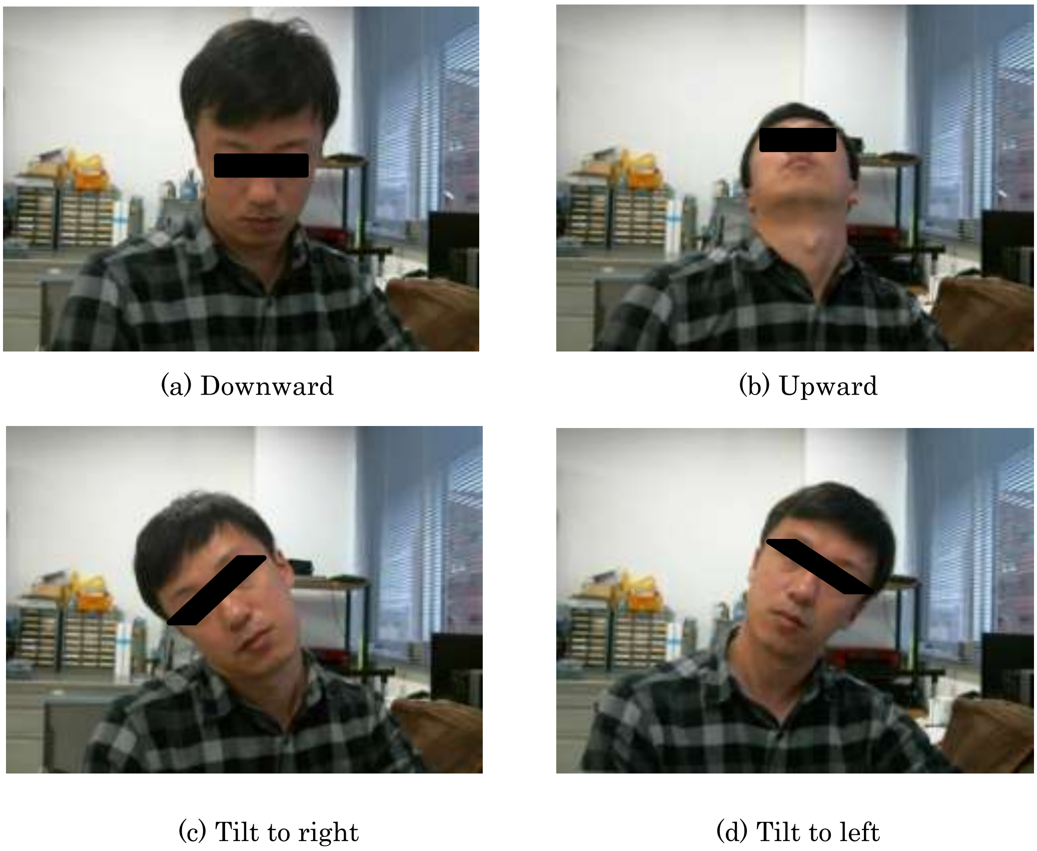


Fig.3.26 position relationship between mouth and nostril

And normally, in meal support device, we scan the front of the human face, and as shown in Fig.3.27 these are the examples of conditions of human face, in these figures, we realize that the center of double nostrils and the mid of the mouth are always in the same line which is the midline of face, so we can locate the area which we can always detect the position of the mouth.



[redacted]

(a) Downward

[redacted]

(b) Upward

[redacted]

(c) Tilt to right

[redacted]

(d) Tilt to left

Fig.3.27 several different condition of face posture

3.8 **The detection of mouth**

Here we propose a new detection method for the mouth position based on the relative position with respect to the nostrils. As shown in Fig.3.28, we consider the perpendicular bisector of two nostrils. The center point between the two nostrils can be described by

$$\begin{cases} y_c = \frac{y_{right} - y_{left}}{2} \\ x_c = \frac{x_{right} - x_{left}}{2} \end{cases} \tag{3.16}$$

Where, (x_{right}, y_{right}) and (x_{left}, y_{left}) are the position of the right and left nostrils respectively. By using this center point and the straight line formula

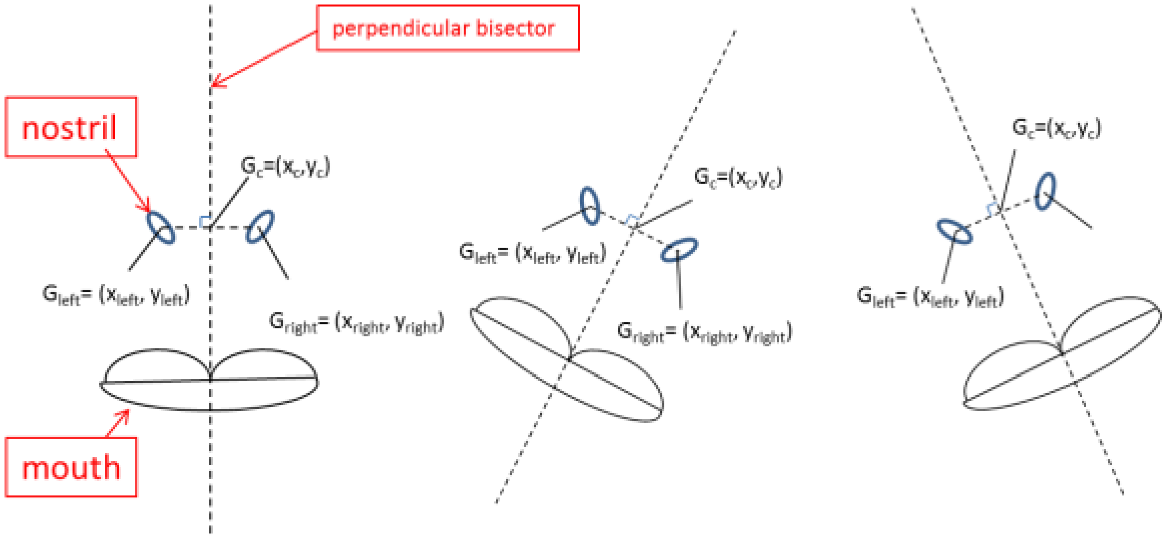
$$y = kx + b \tag{3.17}$$

And in this case

$$\begin{cases} k = -\frac{x_{right} - x_{left}}{y_{right} - y_{left}} \\ b = y_c + x_c \left(\frac{x_{right} - x_{left}}{y_{right} - y_{left}}\right) \end{cases} \tag{3.18}$$

And according to the formula (3.16) the perpendicular bisector can be described by

$$y = -\frac{x_c}{y_c} x + \frac{y_c^2 + x_c^2}{y_c} \tag{3.19}$$


perpendicular bisector
nostril
$G_c=(x_c,y_c)$
$G_{left}=(x_{left}, y_{left})$
$G_{right}=(x_{right}, y_{right})$
mouth
$G_c=(x_c,y_c)$
$G_{left}=(x_{left}, y_{left})$
$G_{right}=(x_{right}, y_{right})$
$G_c=(x_c,y_c)$
$G_{left}=(x_{left}, y_{left})$

(a) front face (b) tilt to left (c) tilt to right

Fig.3.28 different direction of face

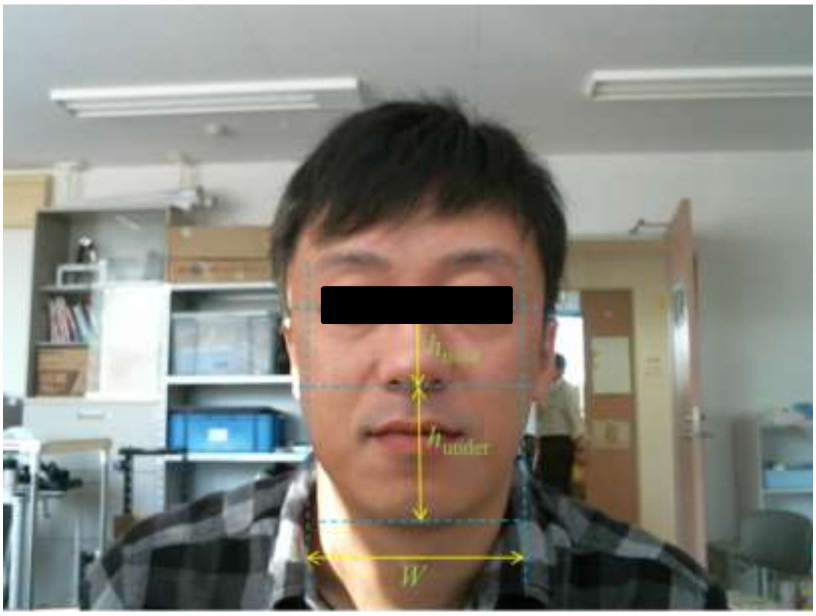

[redacted]

Fig.3.29 Size of human face

In this research, we also focus on the detection of eyes [54], as shown in Fig.3.29 we use the actually size of human face [53] to determine the area which we search of mouth, the average distance from eyes to the nostrils is h_{nose} the average distance between nostrils and chin is h_{under} and the average wide of mandible is W, according to the data of human face size, the ratio between h_{under}, h_{nose} and W is calculated as

$$\begin{cases} \dfrac{W}{h_{nose}} = 2.01 \\ \dfrac{h_{under}}{h_{nose}} = 1.35 \end{cases} \tag{3.20}$$

And in our system, we detect the distance between eyes [54] and nostril as $\tilde{\mathrm{h}}_{nose}$ in real time, and according to the proportion between these h_{nose}, h_{under}, and W, the $\tilde{h}_{under}$ and $\widetilde{W}$ which we use to calculate the area of mouth detection will be work out by the formula below,

$$\begin{cases} \widetilde{W} = \dfrac{W}{h_{nose}} \tilde{h}_{nose} \\ \tilde{h}_{under} = \dfrac{h_{under}}{h_{nose}} \tilde{h}_{nose} \end{cases} \tag{3.21}$$

After we calculate the area of the mouth, we reduce this area by experience below the nostril by α=5[pixel] and β=5[pixel] in W sides as shown in Fig.3.30.

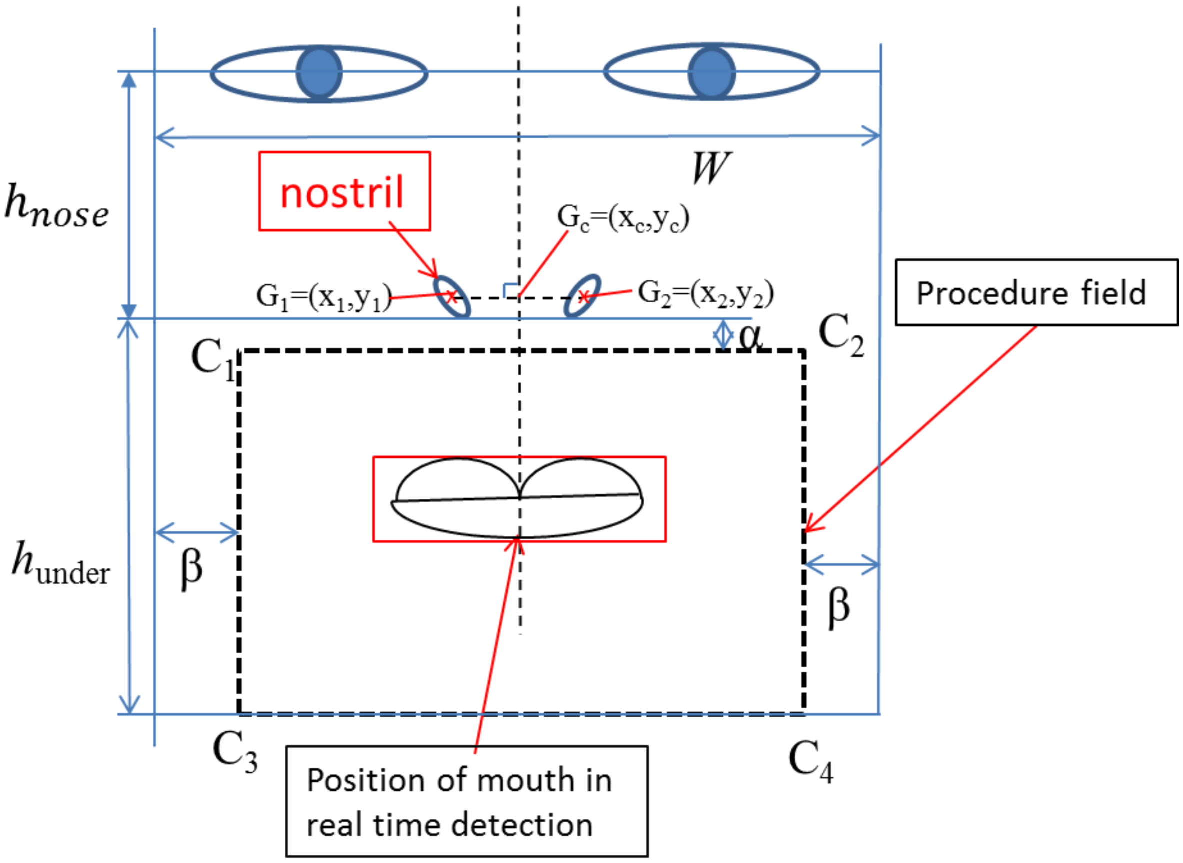


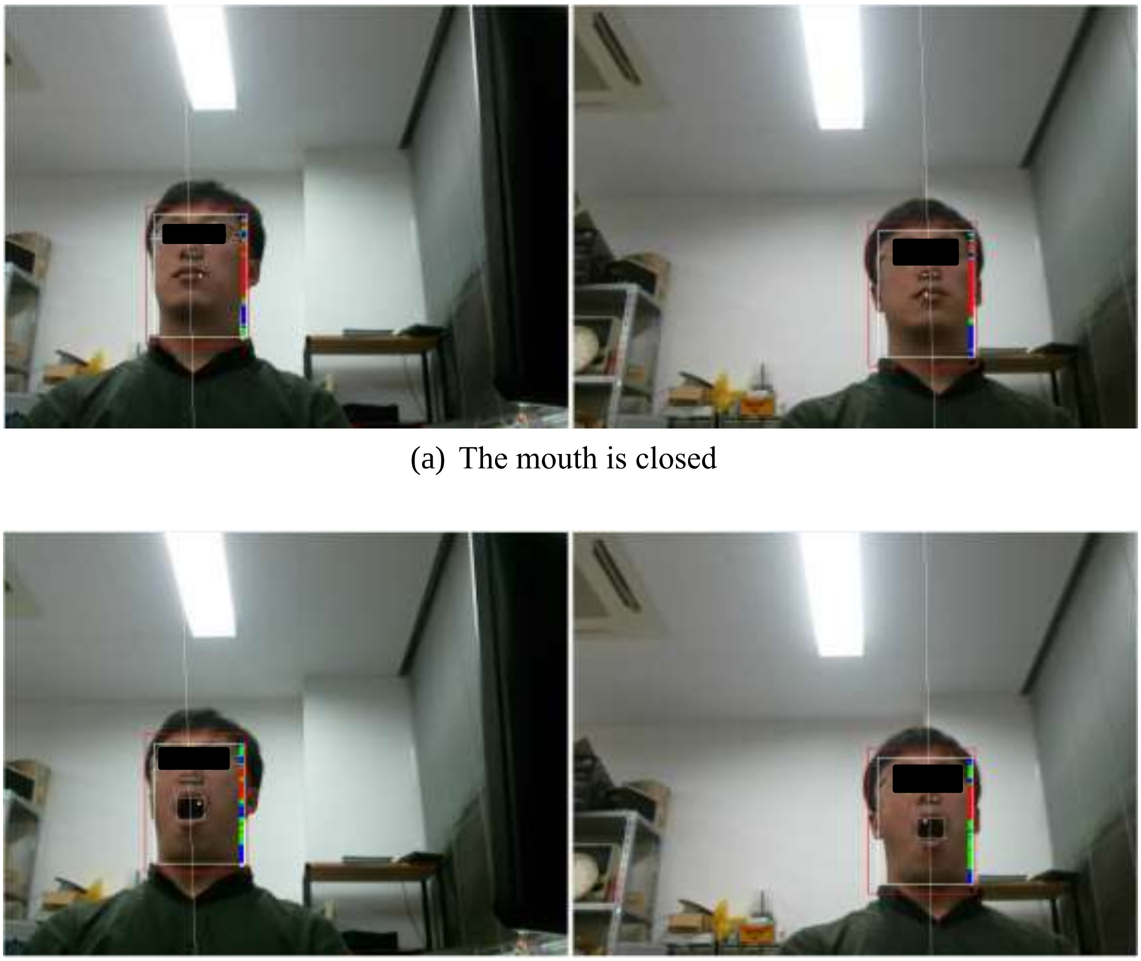
Fig.3.30 detection of mouth area

And the detection area rectangle $C_1C_2C_4C_3$ can be calculated by the formula below

$$\begin{cases} C_1(x_1,y_1) = \left(x_c - \frac{W}{2} + \beta, y_c + \alpha\right) \\ C_2(x_2,y_2) = \left(x_c + \frac{W}{2} - \beta, y_c + \alpha\right) \\ C_3(x_3,y_3) = \left(x_c - \frac{W}{2} + \beta, y_c + h_{under}\right) \\ C_4(x_4,y_4) = \left(x_c + \frac{W}{2} - \beta, y_c + h_{under}\right) \end{cases} \tag{3.22}$$

3.9 **Verification experiment**

In the verification experiment we try to detect the position of the mouth by our constructed system. Fig.3.31 (a) and (b) show the condition when the mouth is closed and opened, respectively. The As shown in Fig.3.31 (a), when the area of the mouth is small than the threshold set, the rectangle around the mouth wasn't displayed, on the other hand, as shown in Fig.3.31 (b) when the area of mouth is larger than the threshold set, the rectangle around the mouth appeared. So the result shows that the system can detect the mouth position correctly, and the condition of the mouth (open or close) can be judged.

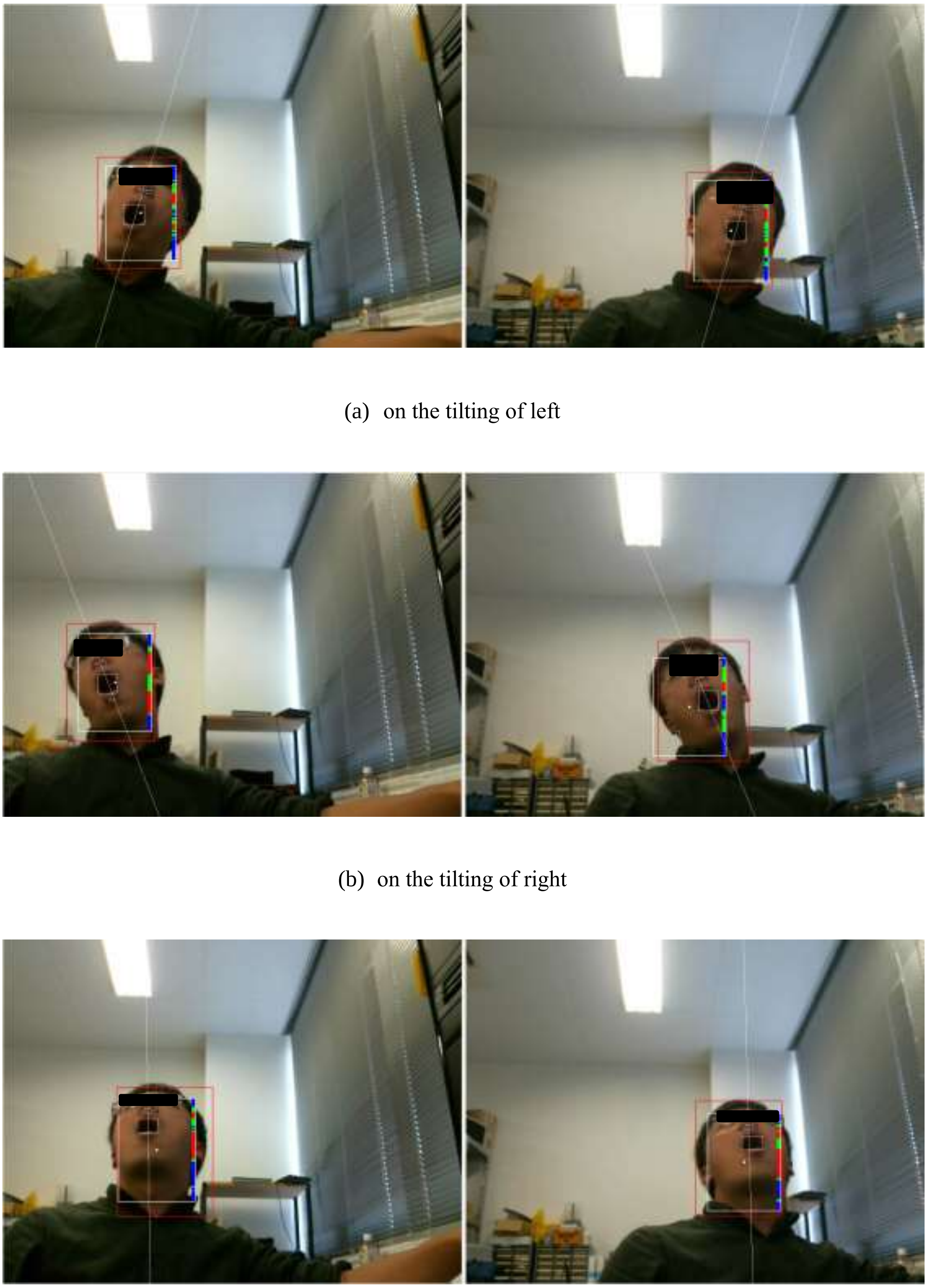


(a) The mouth is closed

(b) The mouth is open

Fig.3.31 result of verification experiment

Next, we consider the condition of face posture, tilting, upward, and head drop. As shown in Fig.3.32 (a) - (c), the system can detect the mouth when the head is tilting to the left and right. If the face is upward, the system can also recognize the mouth position correctly. But when the face is directed to the downward, the system doesn’t display the rectangle as shown in Fig.3.32 (d) because it’s difficult for patient to swallow food when the head direct to the downward, in order to avoid dysphagia, this system does not display mouth position when the body is in inappropriate posture.



[redacted]

[redacted]

(a) on the tilting of left

[redacted]

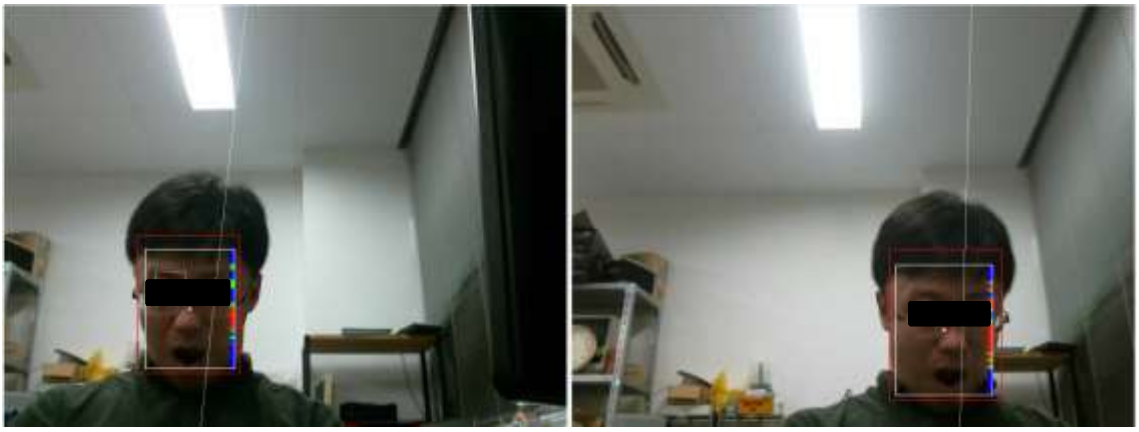
[redacted]

(b) on the tilting of right

[redacted]

[redacted]

(c) on the upward



(d) on the head drop

Fig.3.32 different condition of face posture

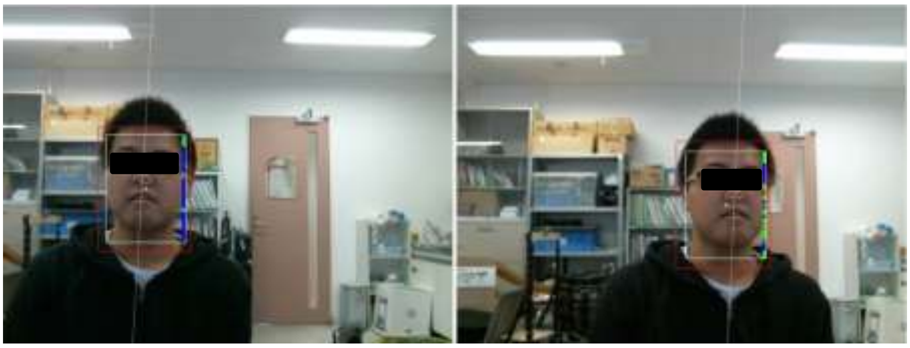
Next, we carried out the experiments by multiple subjects. Twelve subjects (ten men and two women) participated in the experiment. We confirmed the detection of face, nostril and mouth under the condition that the USB camera was located in front of the subject. The subjects were directed to open the mouth. Moreover, we carried out the same experiment in cases where the camera was in the left and right side of the subject. The experiment results are shown in Table.3.3, the circle mark means successful detection, and the crosses mark means that detection was failed. Some example images of the detection are shown in Fig.3.33. As shown in this figure, the color of skin is different among the subject and the brightness of skin is also changed according to the face direction with respect to the celling light. Regardless of this condition, the results of Table.3.3 says that the system could detect all subject faces. As for the nostril, no matter what shape and area the two nostrils were, the nostril of all subjects could also be detected. Furthermore, the mouth open state of all subjects were recognized and the rectangle around the mouth consequently appeared. In the side detection, the mouth could be detected in most conditions, but the mouth of subject “I” couldn’t be detected accurately in two cameras. The subject “I” has the thick moustache shown in Fig3.33(s) and (t), the moustache has deep color and effected the measurement of human face area, and when the subject face to the side, mouth and nostril was outside the area of face, so the mouth could not be detected. However, it is possible to solve this problem by moving the position of the camera to the front of the subject.

From the above results, we considered that the system is enough to detect the mouth position and open/close states and applied it to the manipulation of meal support equipment.

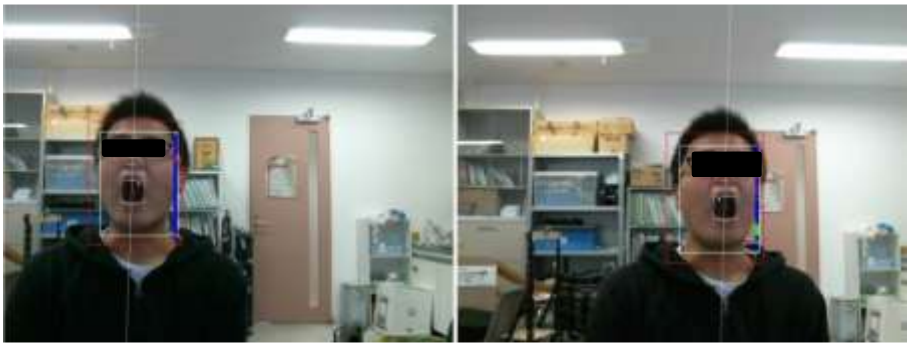
Table 3.3 Experimental examples

	In front side		Left side		Right side	
Subject	Nostril	Mouth	Nostril	Mouth	Nostril	Mouth
A	○	○	○	○	○	○
B	○	○	○	○	○	○
C	○	○	○	○	○	○
D	○	○	○	○	○	○
E	○	○	○	○	○	○
F	○	○	○	○	○	○
G	○	○	○	○	○	○
H	○	○	○	○	○	○
I	○	○	×	×	×	×
J	○	○	○	○	○	○
K	○	○	○	○	○	○
L	○	○	○	○	○	○

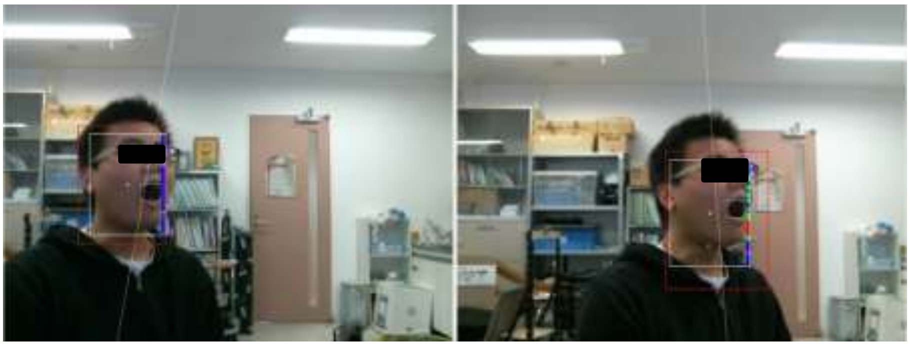
○ Detected △ Misrecognized ×Not detected



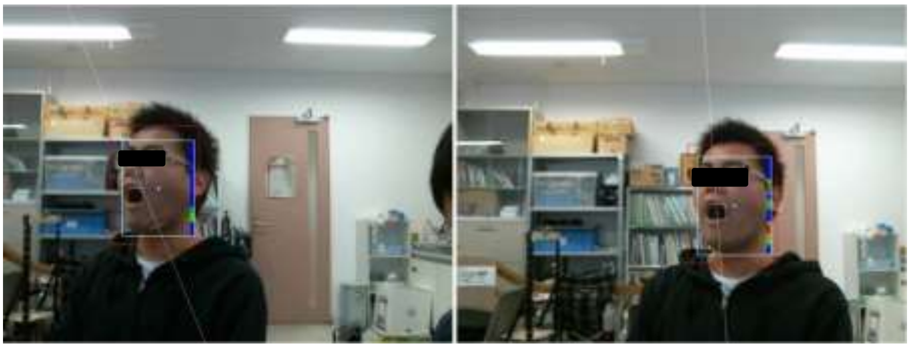
(a) Front detection of subject A



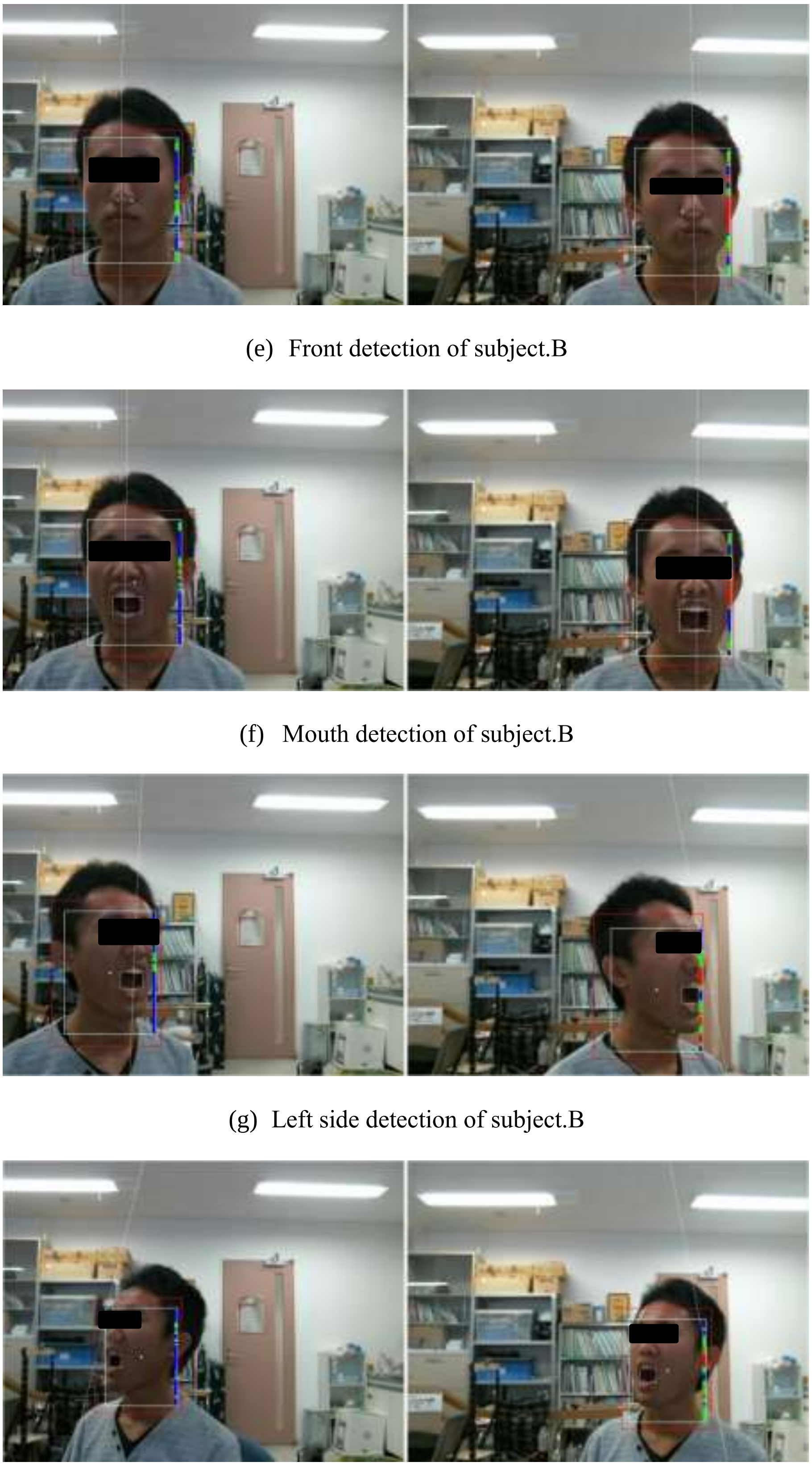
(b) Mouth detection of subject.A



(c) Right side detection of subject.A



(d) Left side detection of subject.A



[redacted] [redacted]

(e) Front detection of subject.B

[redacted] [redacted]

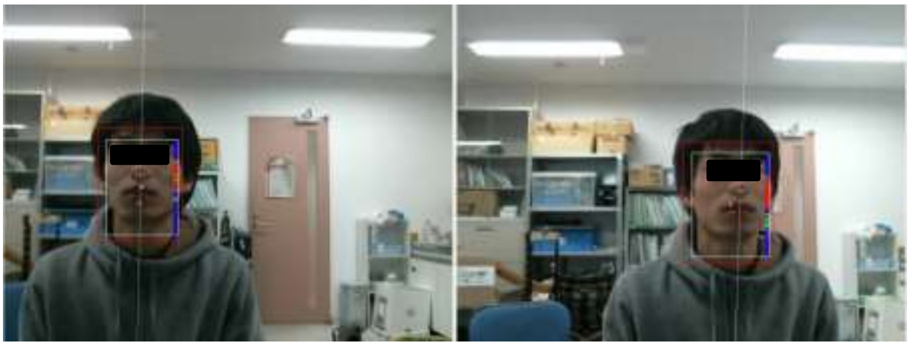
(f) Mouth detection of subject.B

[redacted] [redacted]

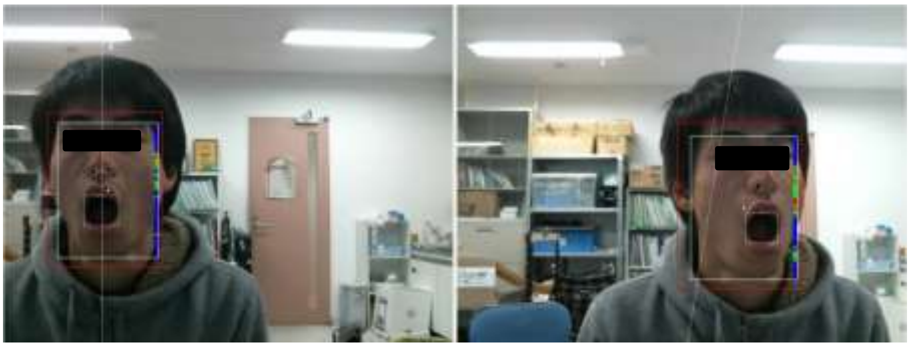
(g) Left side detection of subject.B

[redacted] [redacted]

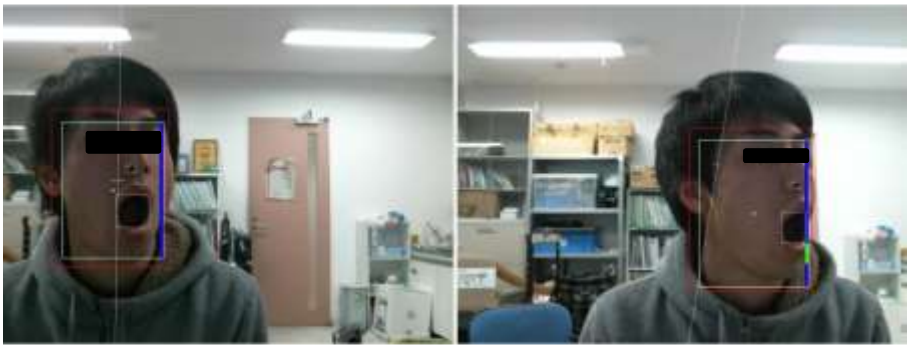
(h) Right side detection of subject.B



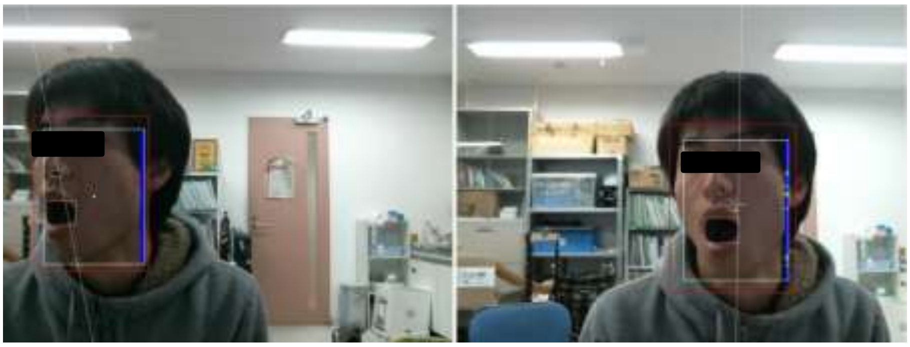
(i) Front detection of subject.C



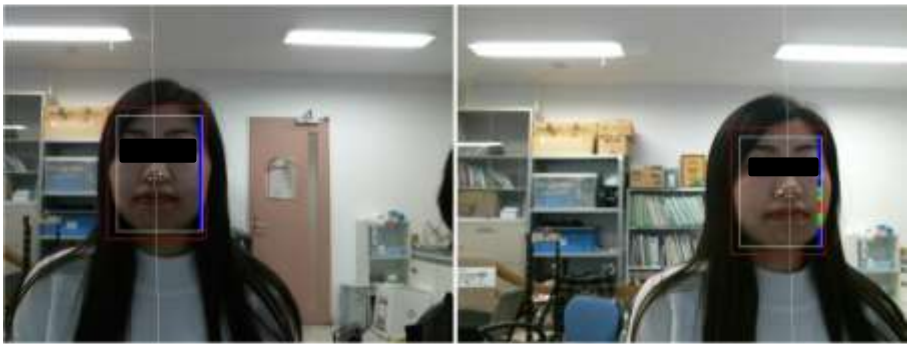
(j) Mouth detection of subject.C



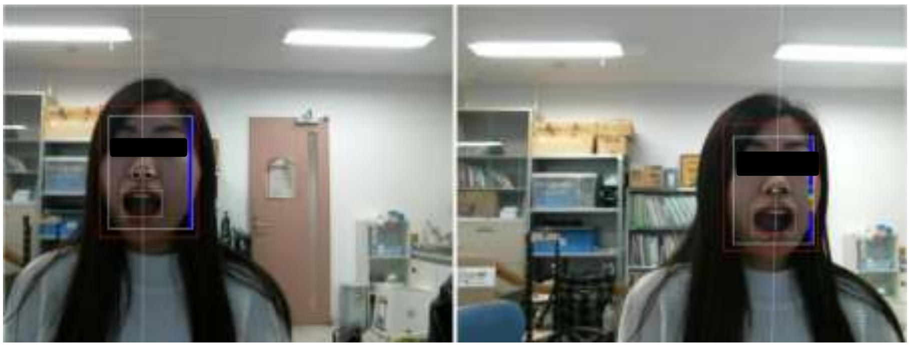
(k) Left detection of subject.C



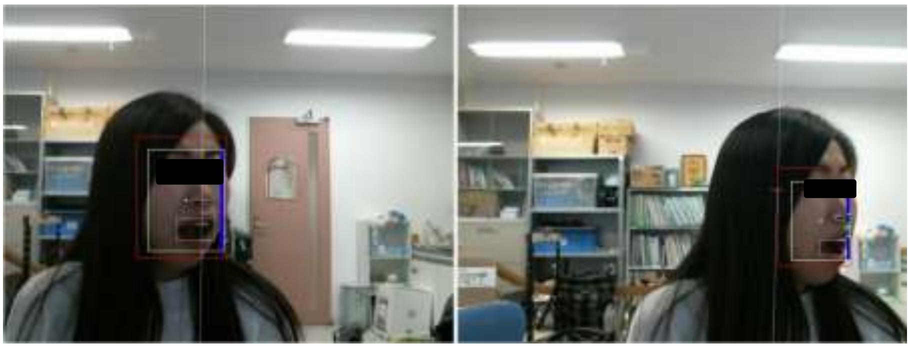
(l) Right detection of subject.C



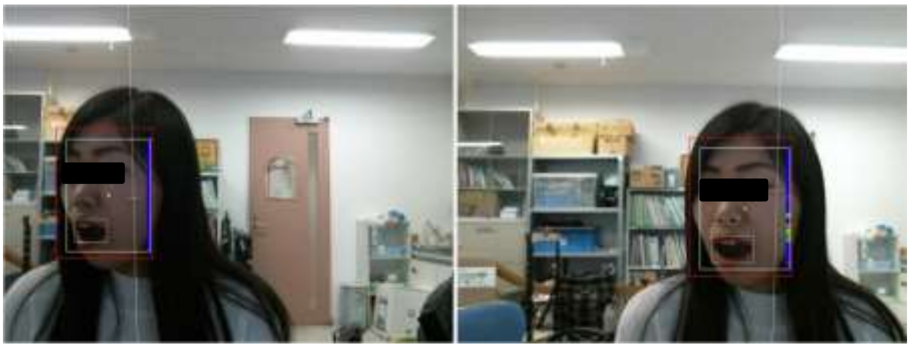
(m) Front detection of subject.K



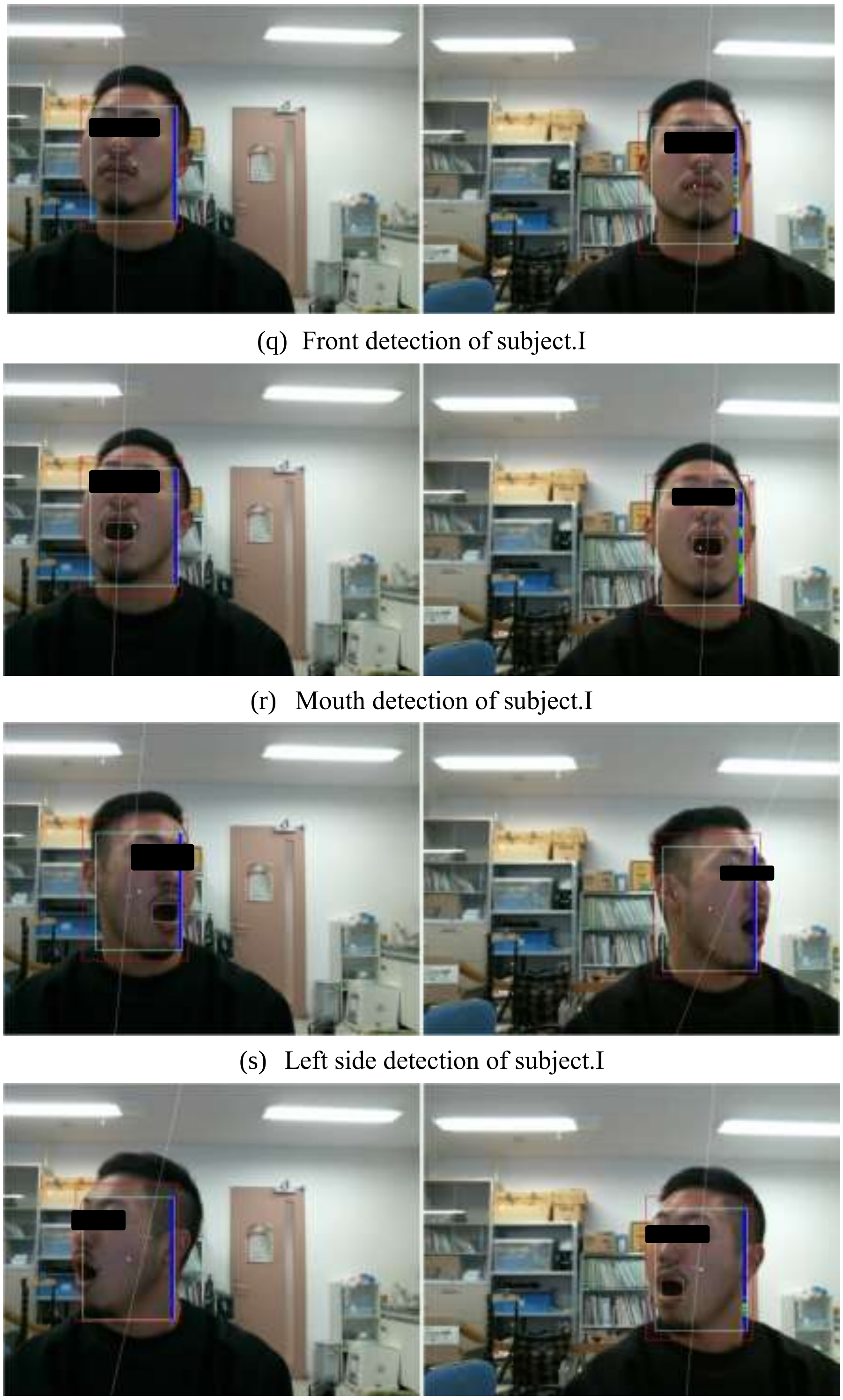
(n) Mouth detection of subject.K



(o) Left side detection of subject.K



(p) Right side detection of subject.K



(q) Front detection of subject.I

(r) Mouth detection of subject.I

(s) Left side detection of subject.I

(t) Right side detection of subject.I

Fig.3.33 Result example of face detection

Chapter 4 Application to Manipulator Control

4.1 **Manipulator control**

4.1.1 **Kinematics and inverse kinematics of manipulators**

In this system, we combined a three-link manipulator and a three-dimensional measurement system with stereo camera and operating the manipulator with the coordinates (X, Y, Z) in the measured three-dimensional space.

The positional relationship between the measurement system and the manipulator is shown in Fig.4.1， The plane *x-y* of the manipulator and the plane $X - Z$ of the measurement system were made into coplanar, and the axis *y* and the axis *Y* were aligned, and the axis *x* and the axis X were arranged in parallel.

As shown in Fig.4.1, the conversion from the coordinates (X, Y, Z) in the three-dimensional space to the tip of target position (x_3, y_3) is expressed by the following equation.

$$\begin{cases} \mathrm{x_3} = -\mathrm{X} \\ \mathrm{y_3} = \mathrm{Z} - 350 \end{cases} \tag{4.1}$$

By assigned the coordinates calculated by this transformation, we control the manipulator and verify the operation.

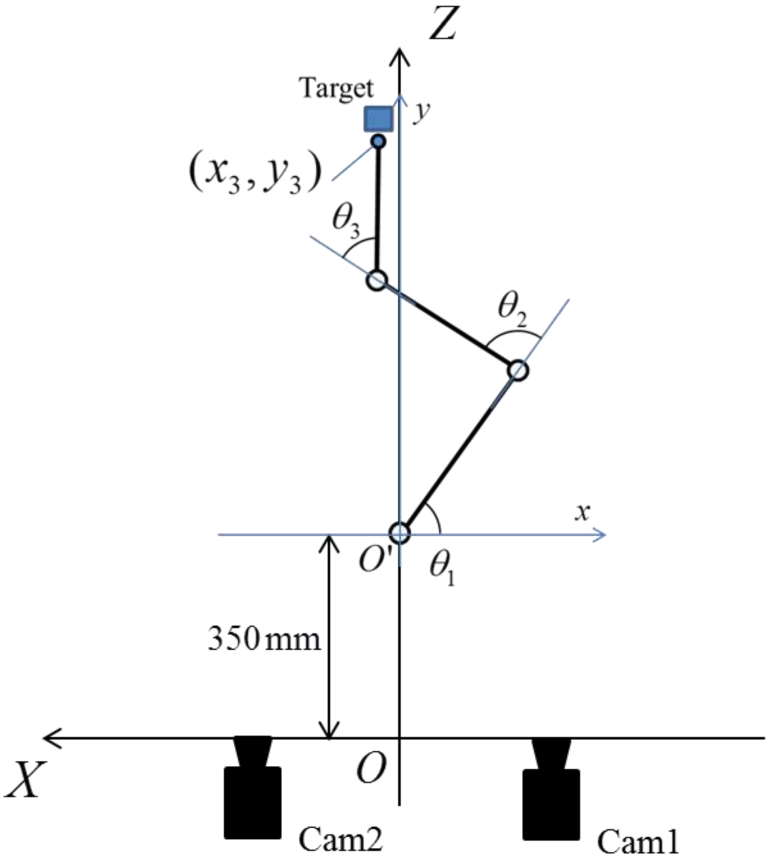


Fig.4.1 three-link manipulator with stereo camera system

Fig.4.2 shows the outline of the manipulator. In this research, we use three-link manipulator. For each joint, we use the RC servomotor [43] (KONDO KRS-6003HV: maximum operating angle 270 [°], torque 67.0 [kg · cm], maximum speed 0.22 [sec / 60 °]) as shown in Fig. 4.3. Multiple RC servomotors can be controlled in real time via serial communication via the control board (KONDO RCB 4 - HV: M16C / 26 A manufactured by Renesas Technology Corporation) shown in Fig.4.4. The coordinate system of manipulator is shown in Fig.4.5, the positions (x_1, y_1), (x_2, y_2) and (x_3, y_3) of each joint can be expressed by the following equations.

$$\begin{cases} x_1 = l_1 \cos(\theta_1) \\ y_1 = l_1 \sin(\theta_1) \end{cases} \tag{4.2}$$

$$\begin{cases} x_2 = x_1 + l_2 \cos(\theta_1 + \theta_2) \\ y_2 = y_1 + l_2 \sin(\theta_1 + \theta_2) \end{cases} \tag{4.3}$$

$$\begin{cases} x_3 = x_2 + l_3 \cos(\theta_1 + \theta_2 + \theta_3) \\ y_3 = y_2 + l_3 \sin(\theta_1 + \theta_2 + \theta_3) \end{cases} \tag{4.4}$$

where θ_1, θ_2, θ_3 are the angles of the joints. The link length of the manipulator is $l_1 = 69$ [mm], $l_2 = 66$ [mm], $l_3 = 85$ [mm].

Let us consider a state in which the third link is always parallel to the *y* axis as shown in Fig.4.6. In this case, the joint angles $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_3$ are assigned by the following conditions.

$$\hat{\theta}_1 + \hat{\theta}_2 + \hat{\theta}_3 = \frac{\pi}{2} \tag{4.5}$$

If the target position of the tip is$(x_3, y_3) = (\hat{x}_3, \hat{y}_3)$, the position $(\hat{x}_2, \hat{y}_2)$of the second joint is obtained from the expression (4.4).

$$\begin{cases} \hat{x}_2 = \hat{x}_3 - l_3 \cos(\hat{\theta}_1 + \hat{\theta}_2 + \hat{\theta}_3) = \hat{x}_3 - l_3 \cos(\frac{\pi}{2}) = \hat{x}_3 \\ \hat{y}_2 = \hat{y}_3 - l_3 \sin(\hat{\theta}_1 + \hat{\theta}_2 + \hat{\theta}_3) = \hat{y}_3 - l_3 \sin(\frac{\pi}{2}) = \hat{y}_3 - l_3 \end{cases} \tag{4.6}$$

$$\hat{\theta}_2 = \cos^{-1}\left(\frac{\hat{x}_2{}^2 + \hat{y}_2{}^2 - l_1{}^2 - l_2{}^2}{2 l_1 l_2} \right) \tag{4.7}$$

$$\hat{\theta}_1 = \sin^{-1}\left(\frac{\hat{y}_2}{\sqrt{\hat{x}_2{}^2 + \hat{y}_2{}^2}} \right) - \phi \tag{4.8}$$

where

$$\phi = \tan^{-1}\left(\frac{l_2 \sin(\gamma)}{l_1 + l_2 \cos(\gamma)} \right) \tag{4.9}$$

$$\gamma = \begin{cases} \hat{\theta}_2 & (\hat{x}_2 \geq 0) \\ -\hat{\theta}_2 & (\hat{x}_2 < 0) \end{cases} \tag{4.10}$$

From the formulas (4.5), (4.7) and (4.8)

$$\hat{\theta}_3 = \frac{\pi}{2} - \hat{\theta}_1 - \hat{\theta}_2 \tag{4.11}$$

Each joint angle $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_3$ is obtained from the target position (x_3, y_3) namely $(\hat{x}_3, \hat{y}_3)$. In this research, considering about the contact direction to facial feature points, the state in Fig. 4.6 is taken as the basic posture.

4.1.2 RC servo motor angular control command

The angle control of the RC servomotor performs KONDO's original ICS (Interactive Communication System ver. 3.0) command. Values that can be specified as potentio data of the servomotor are 3500 to 11500. Angle control is performed by transmitting this value to the RC servo motor. Since the movable angle of the RC servo motor (KONDO KRS - 6003 HV) is 3 π / 2 [rad]. The angle of 0 [rad] is transmitted when the potentio data of the servomotor is set to 3500. In the case of 3 π / 2 [rad], the potentio data of the servomotor is set to 11500, and the joint angle is determined. At power-on, the joint angle is π / 2 [rad], the potentio data of the servomotor is set to 7500. The value κ of the servo motor's potentio data per unit angle is as follows.

$$\kappa = \frac{11500 - 3500}{\frac{3}{2}\pi} \approx 1697.65 \tag{4.12}$$

Here, the conversion formulas of the angles θ_1, θ_2, θ_3 of each joint and the potentio data Θ_1, Θ_2, Θ_3 of the servo motor are as follows.

$$\begin{cases} \Theta_1 = \left(\dfrac{\pi}{2} - \theta_1\right)\dfrac{1}{\kappa} + 7500 \\ \Theta_2 = -\dfrac{1}{\kappa}\theta_2 + 7500 \\ \Theta_3 = 7500 + \dfrac{1}{\kappa}\theta_3 \end{cases} \tag{4.13}$$

However, κ is the calibration value of equation (4.12).

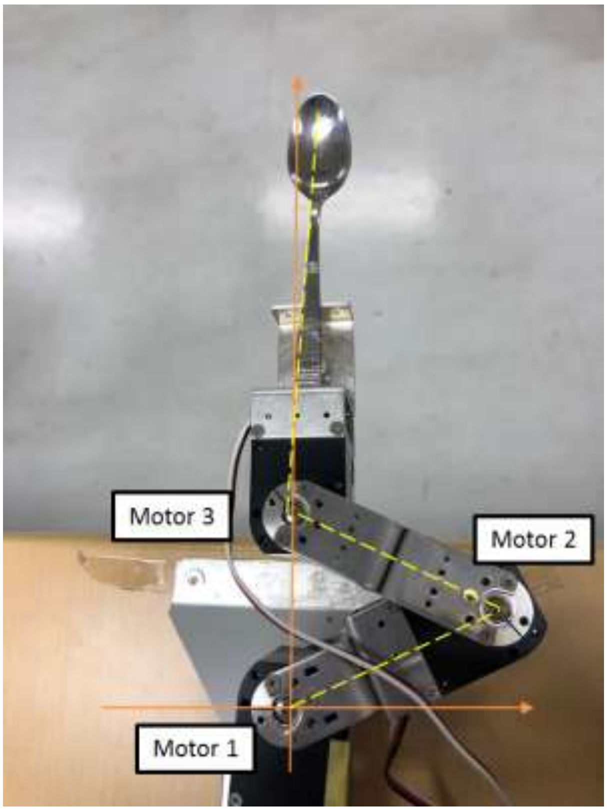


Fig.4.2 three-link manipulator

Fig.4.3 RC servomotor (Kondo KRS-6003HV)

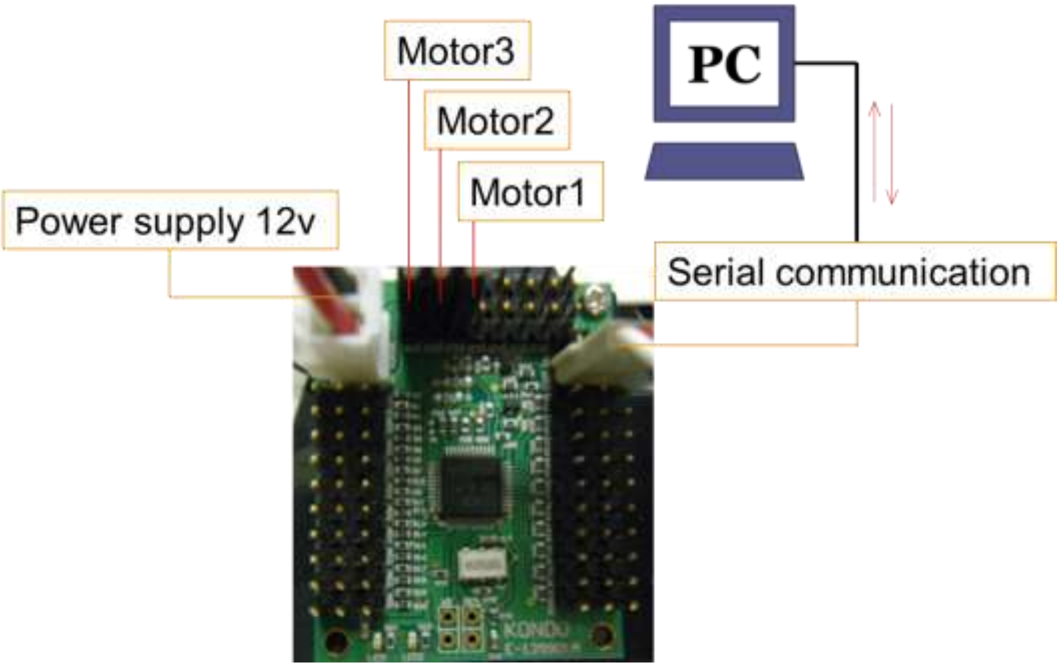


Fig.4.4 Control board (Kondo RCB4-HV)

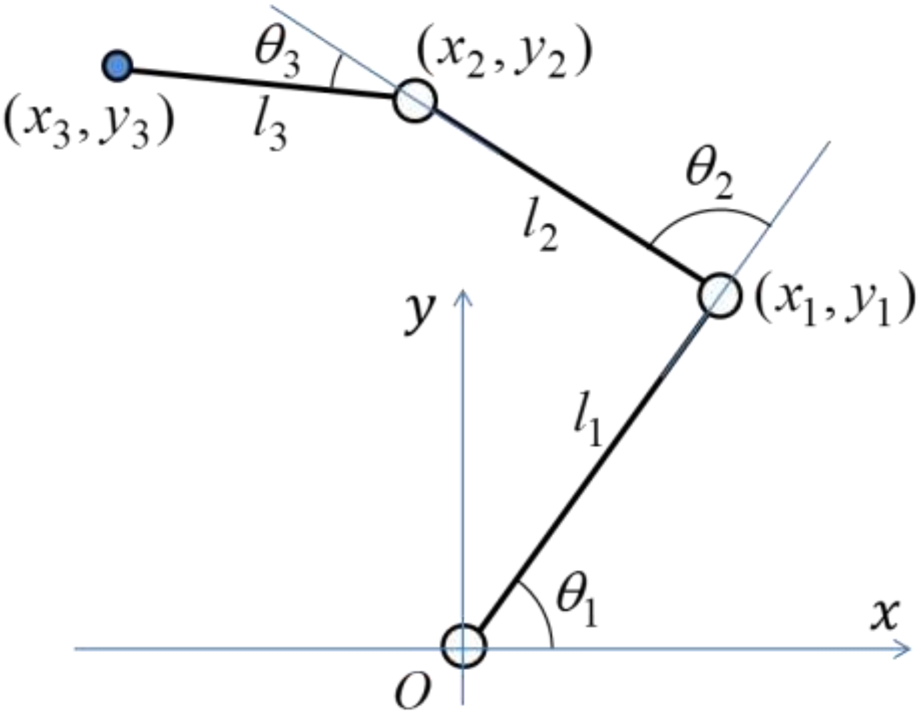


Fig.4.5 three-link manipulator model

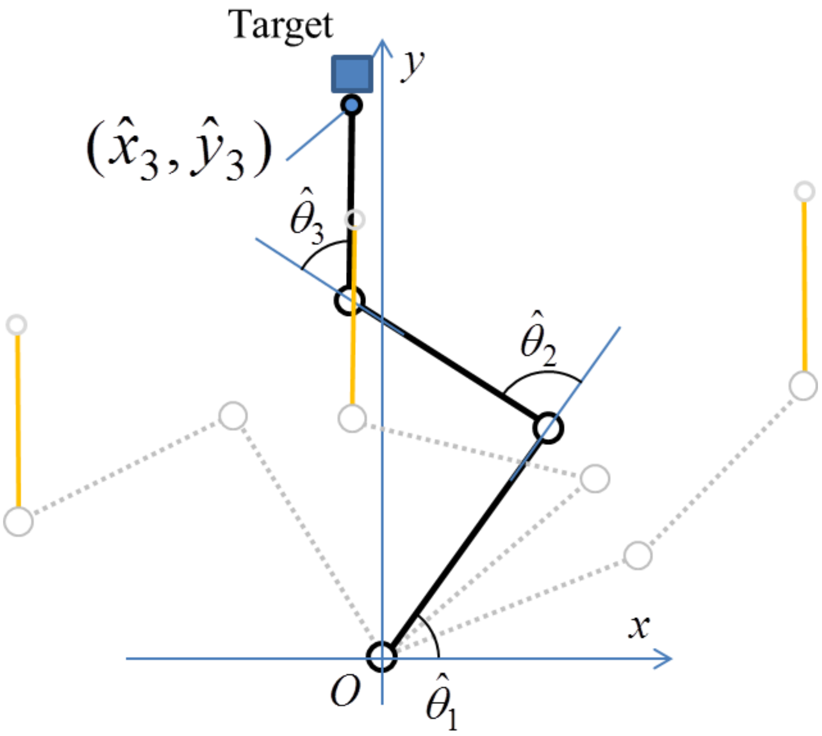


Fig.4.6 Basic posture

4.2 **Verification experiment**

4.2.1 **Experiment method**

In this experiment, we confirm the usefulness of this system by combining a three-link manipulator and a three-dimensional measurement system with stereo camera and operating the manipulator with the coordinates (X, Y, Z) in the measured three-dimensional space. We verified whether the manipulator could touch the correct position, and whether manipulator could carry spoon into human mouth.

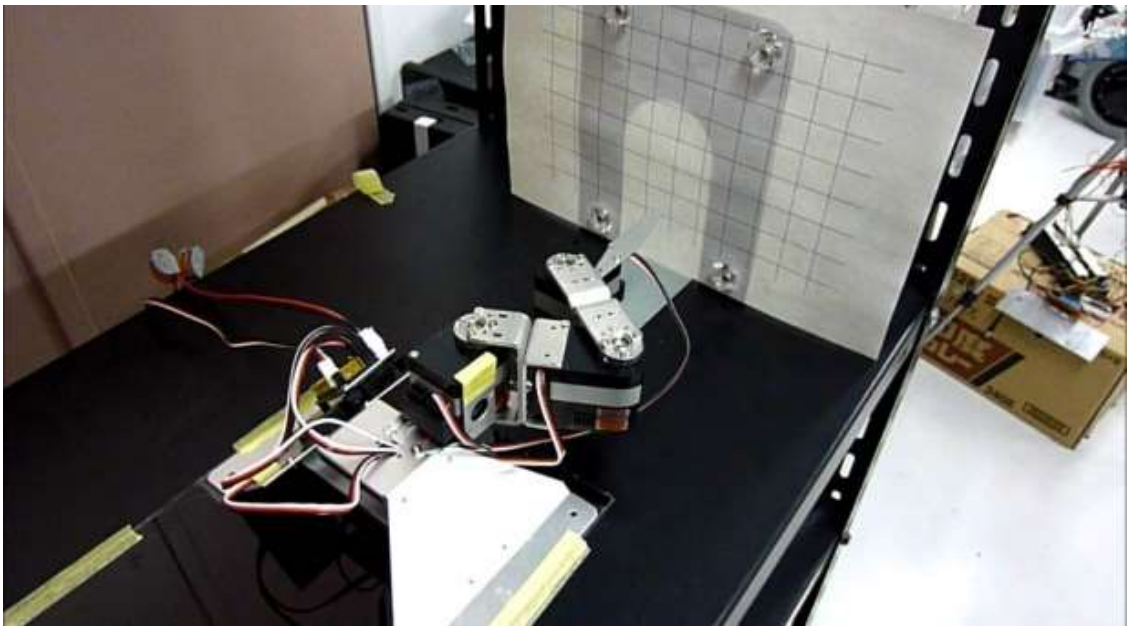
4.2.2 **Experimental results**

Operation experiment of the manipulator is shown in Fig.4.7 and Fig.4.8. These are photographs of the state of the manipulator for each elapsed time $t = 0[\mathrm{s}]$ after starting the operation.

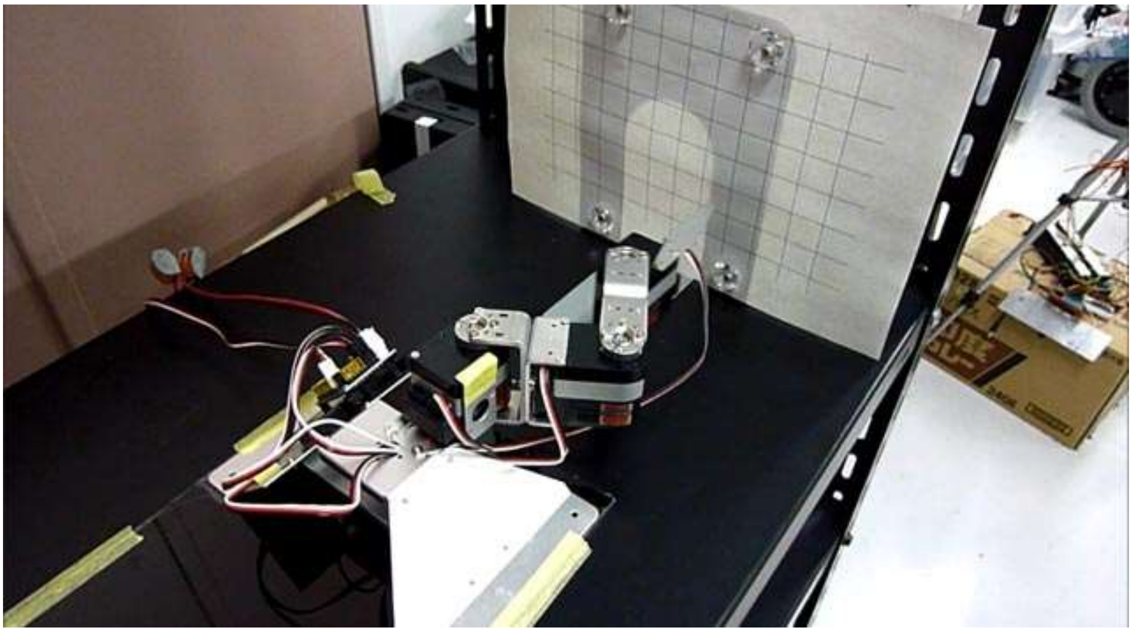
Fig.4.7 shows a lattice-like image, and manipulators are positioned using the coordinates of the intersection as input information. It is possible to confirm that the tip of the arm of the manipulator has reached the target position accurately at the speed of operation time 0.8 [s] by detecting the target intersection point coordinates.

Fig.4.8 shows the input information of the coordinates of the nostrils and mouth detected by the algorithm described in Chapter 3. As shown in this figure, the spoon is put on the tip of the three-link arm. It can be seen that the spoon approaches to the target position, the user’s mouth. In the developed system, the spoon moves to inside the mouse and the user doesn’t need to move the head, due to precision positioning motion. As for the existed device [5], the user must move the head to catch the food on the spoon. The state of positioning at the center of the face can be confirmed and the precision which can be sufficiently used for meal support system is obtained.

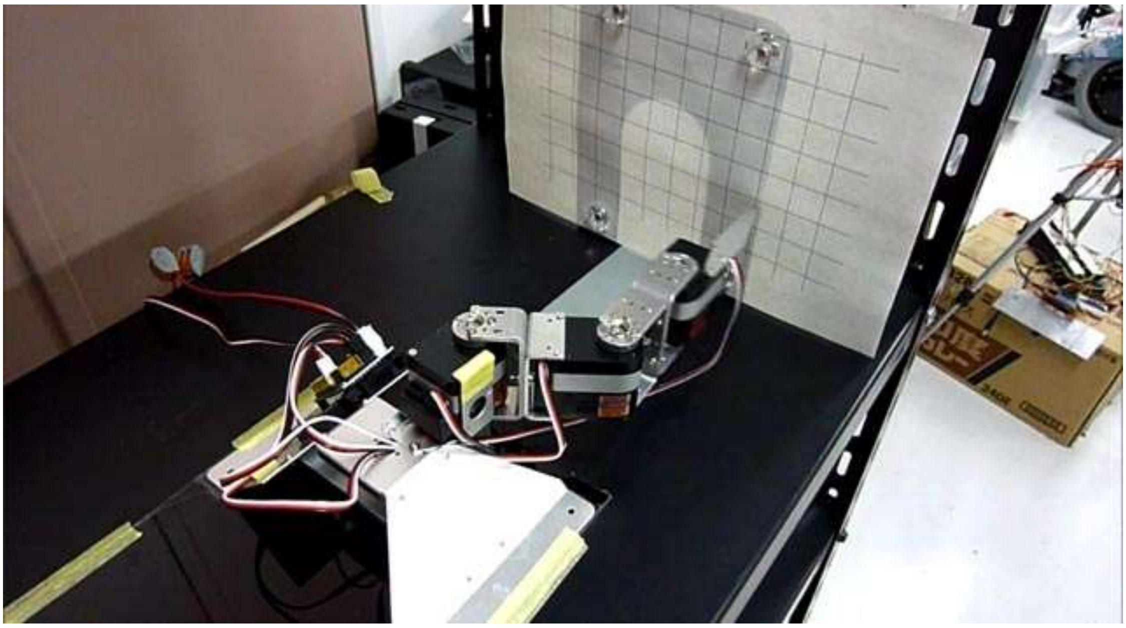
Fig.4.9 shows the scene of the mouth detection. We can see that the rectangle around the user’s mouth appears and the system can recognize it as a target position. When the mouth is closed, the rectangle disappears and the movement motion has been finished. In this figure, the stereo camera was located behind the manipulator, but the user can also select voluntary position of the stereo camera in the developed system. This is one of the merits of the developed system. Fig. 4.10 shows the result of detection when the stereo camera was located at the side of the user. Even though the stereo camera location is changed, the nostril and mouth can be successfully detected to provide the target position to the manipulator.



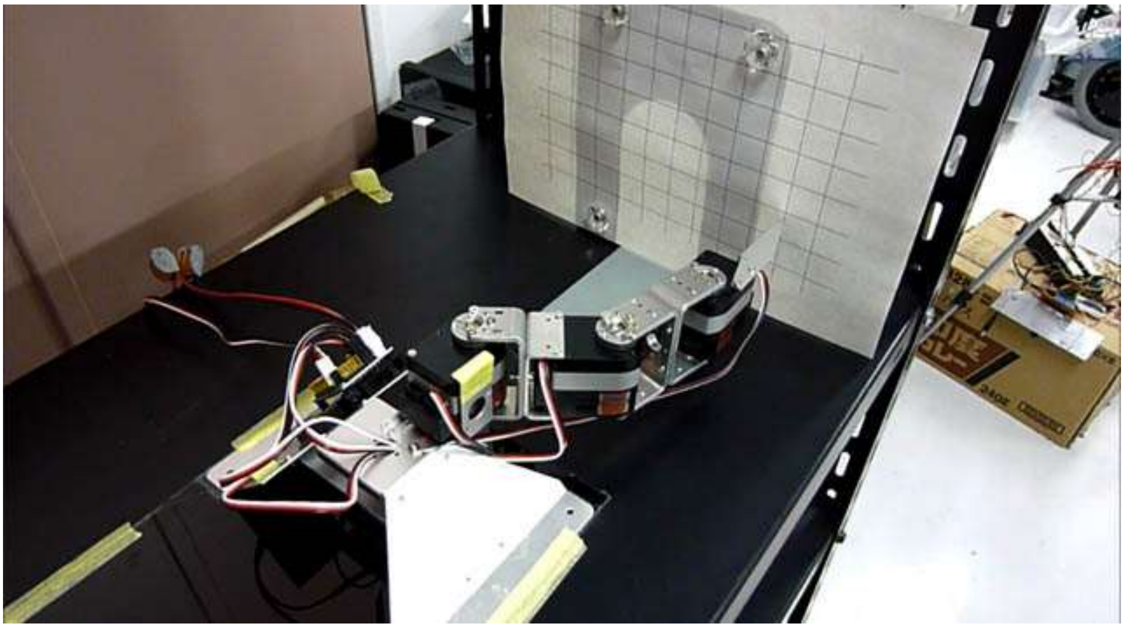
(a) t=0[s]



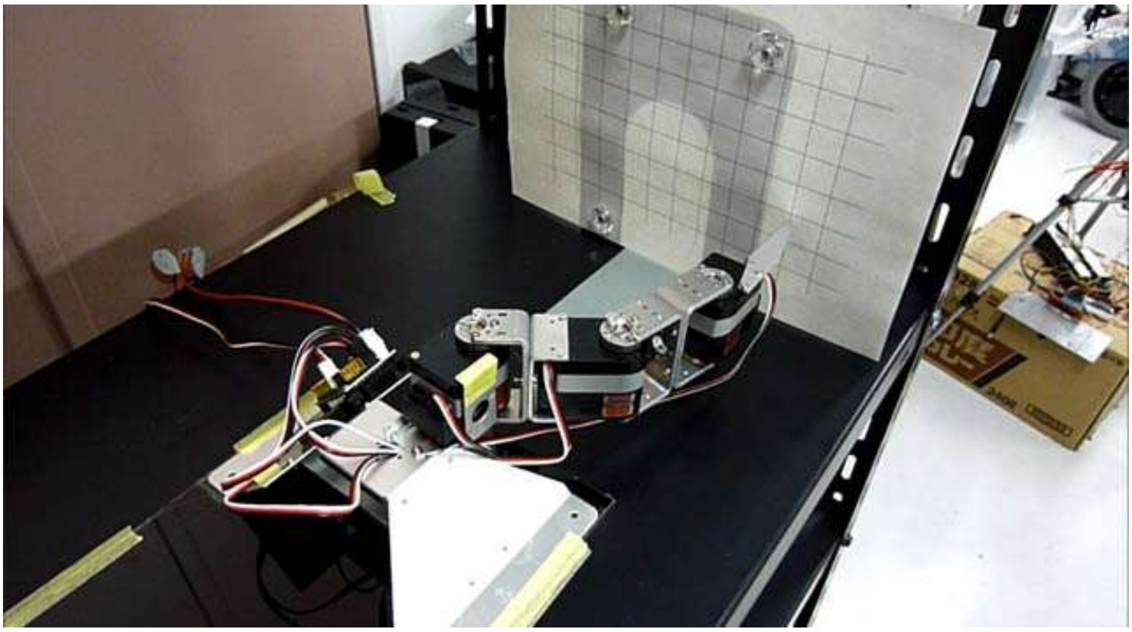
(b) t=0.2[s]



(c) t=0.4[s]

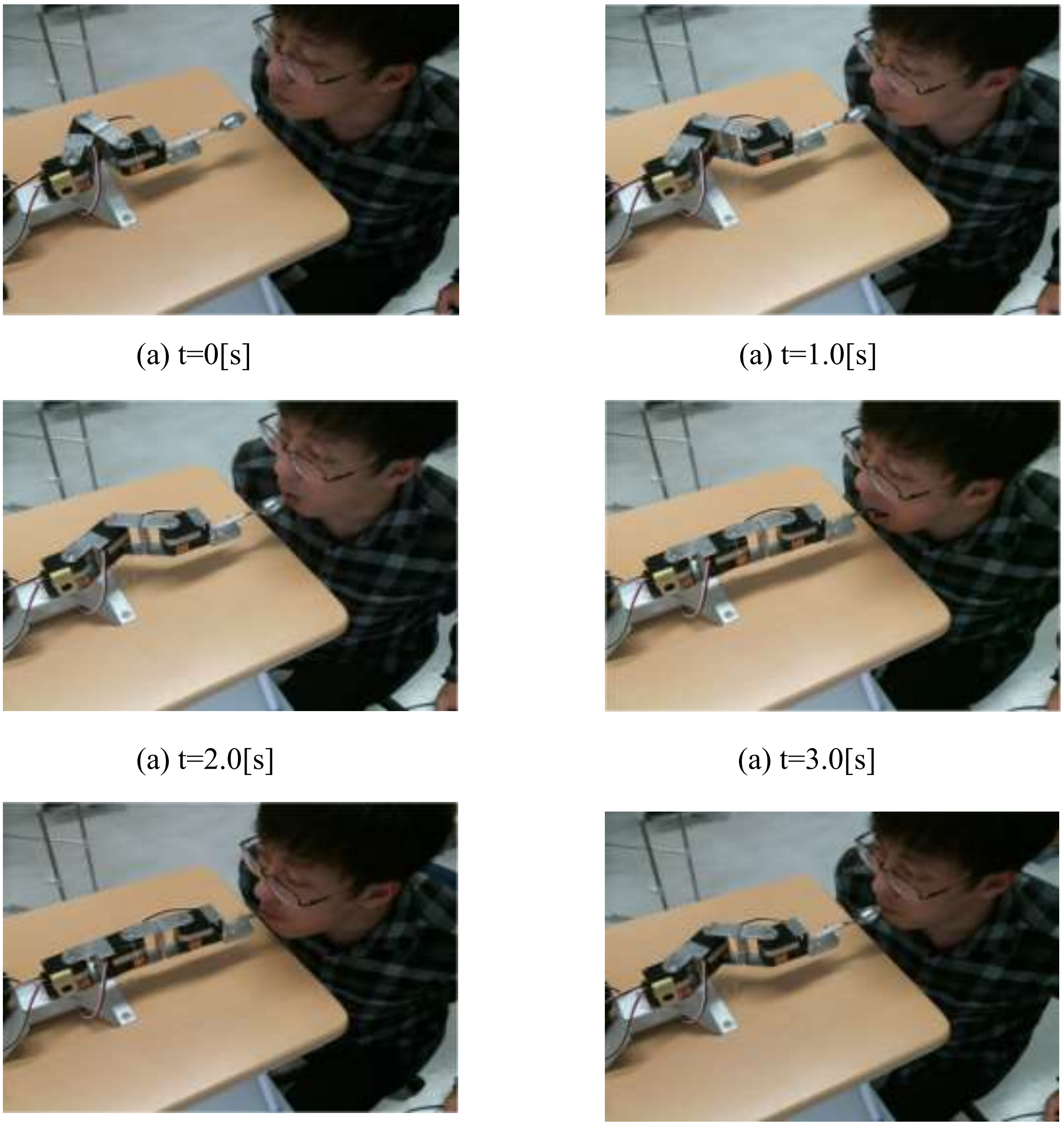


(d) t=0.6[s]



(e) t=0.8[s]

Fig.4.7 Experimental result (Grid pattern)



(a) t=0[s]

(a) t=1.0[s]

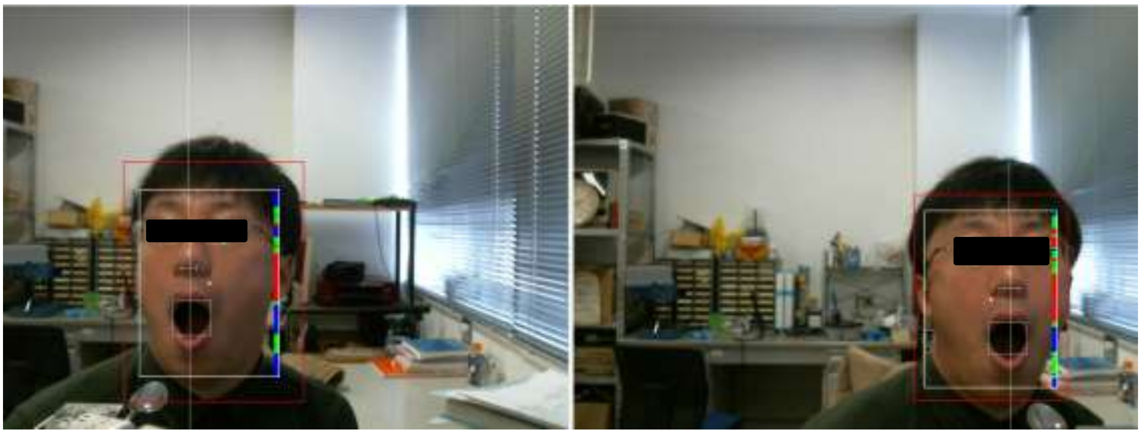
(a) t=2.0[s]

(a) t=3.0[s]

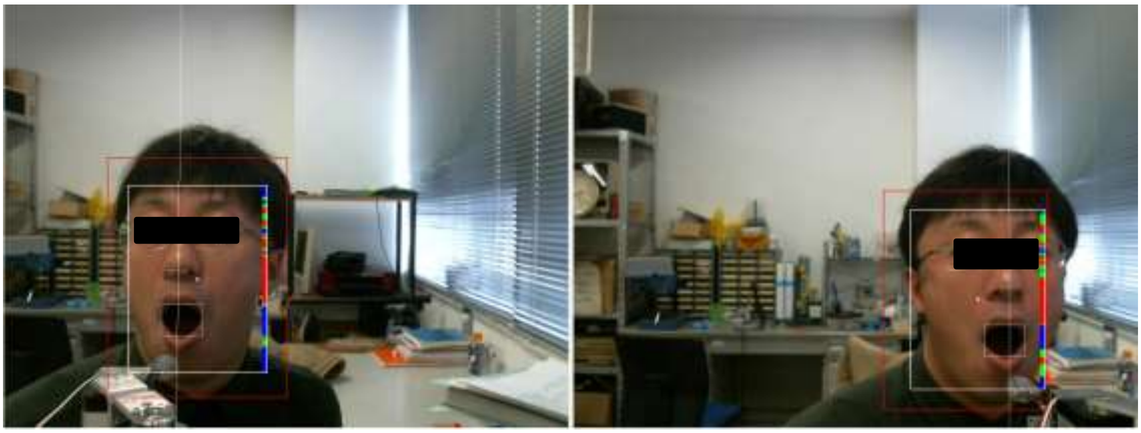
(a) t=4.0[s]

(a) t=5.0[s]

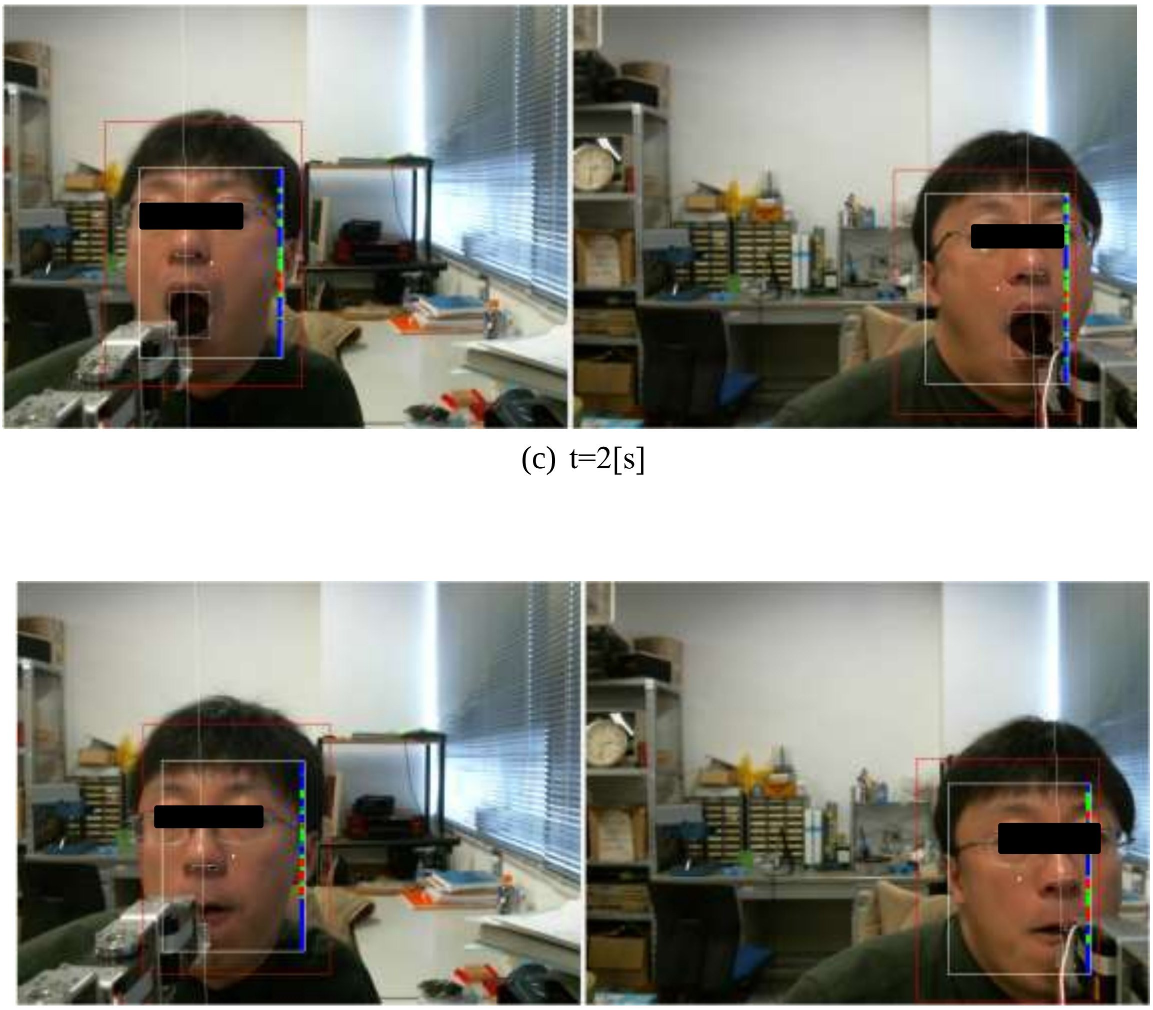
Fig.4. 8 Experimental result (Face)



(a) t=0[s]



(b) t=1[s]



[redacted] [redacted]

(c) t=2[s]

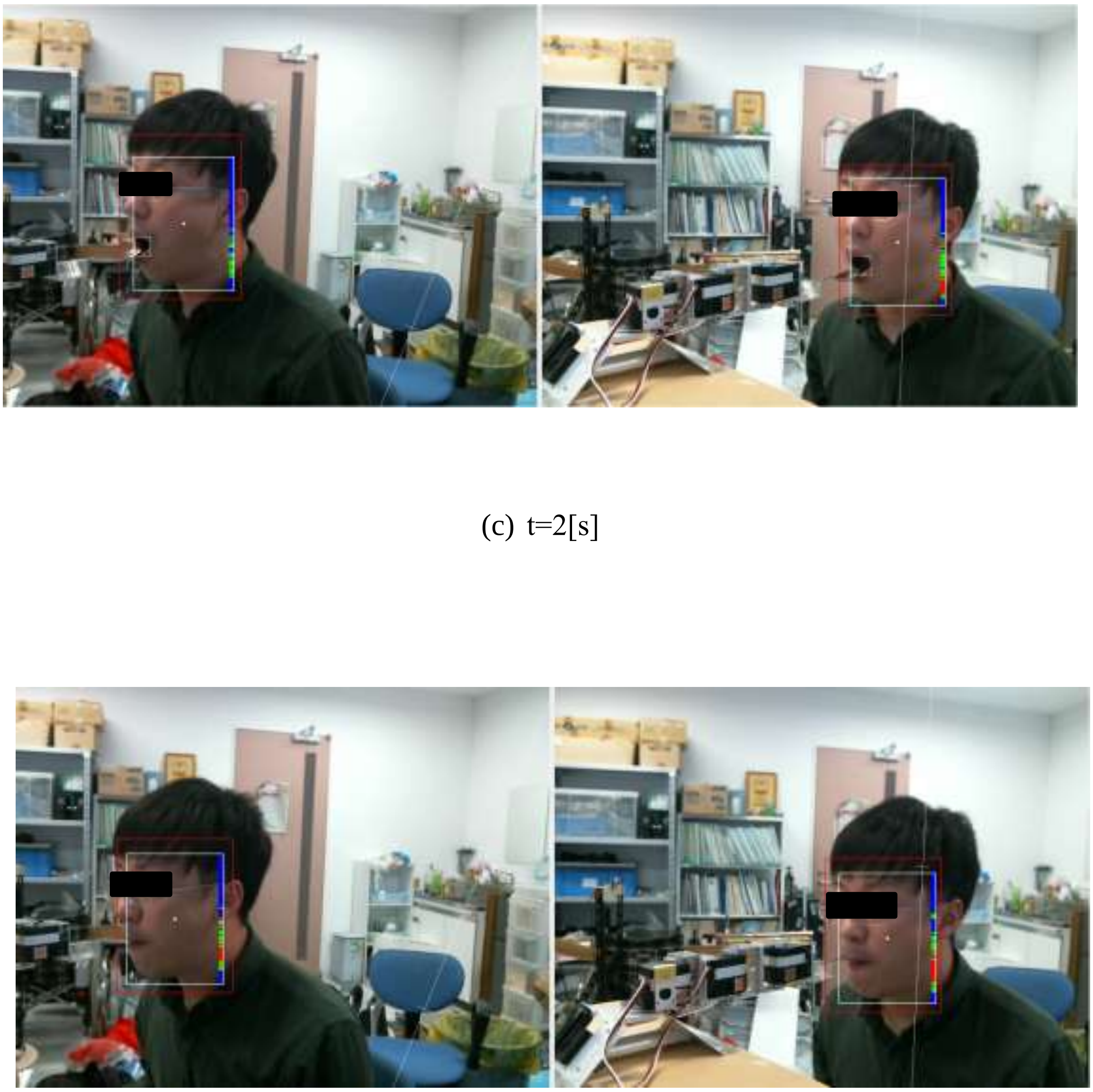
[redacted] [redacted]

(d) mouth is closed

Fig 4.9 Result of detection in front

(a) t=0[s]

(b) t=1[s]



(c) t=2[s]

(d) mouth is closed

Fig.4.10 Result of detection in the side

Chapter 5 Conclusion

In this doctoral thesis, we propose a method for detecting facial feature points in three-dimensional space on the premise of application to meal support equipment for the purpose of supporting disabled persons and elderly people.

Chapter 1 describes the background of the research, existing research, and the significance of this research. Currently, in Japan, the declining birthrate and the aging of society are advancing, the shortage of carers are concerned, and solving the problem of the shortage of workers is one important issue. Especially, assistance of meal, which is one of the pleasure of care receivers, depends largely on the caregiver. In the place such as nursing homes, it always happens that a small number of carers are assisting a large number of meal support. From such a social background, the development of a meal support manipulator has been carried out up to now. However, the main role of the meal support manipulator is to carry food to the mouth, and the users need to move their face in order to put the food into the oral cavity by themselves. Therefore, there is a possibility of giving a burden to the users. In this paper, we propose a method to detect the position of facial feature points in three dimensional space and aim to apply it to meal support manipulator.

In chapter 2, we described the outline of the developed system and the detection of the three-dimensional position by the stereo camera. In this system, two cameras were arranged side by side, and three-dimensional position measurement of the point of object was performed by using parallax from the obtained two images. As a result of measuring the angle of view and optical axis, distance between two cameras, and further calibrating of the image distortion, it showed that we could measure the position of the target with an error within ± 1[%] in the three dimensional space.

In chapter 3, we described the method of automatic detection of face, nostrils and mouth and the results of verification experiments. In order to extract the face of a person from the obtained image, the skin color information in the RGB color space was examined for features under different bright environments. As a result, regardless of the brightness, the component of R was found to be stronger in skin, nostrils and lips than in other areas, and the difference between R and B components was very large. According to these features, formulation was made focusing on the relationship between values, difference, and ratio of RGB, and we found the suitable conditions for skin color. On the other hand, when detecting a face by image processing, it is important to recognize the whole face as one connected component, but it is often not recognized as a connected component depending on the state of the face such as a shadow or eyeglasses. Therefore, rough information of the original image was used, and mosaic processing was applied to the image. For an obtained image of 640 × 480 [pixels], 16 × 16 [pixels] were translated into one block and obtained image was divided into 40 × 30 blocks, and the representative color of each block was obtained from the average color of the representative points. Blocks whose representative colors satisfy the condition of the skin color were extracted and smoothed further to estimate the rough face area. In the estimated face region, binarization and labeling processing were performed on the original image, the nostril candidates were narrowed down and nostrils were detected from the aspect ratio and area of each connected component. On the other hand, note that there were many dark areas around the eyes, mouth, and jaws, but the area around the nostrils has a high proportion of bright skin, and the area around the nostrils is estimated by finding the proportion of skin color in the face area. This characteristic was used to confirm the detected nostril position, and it could prevent misdetection of nostril positions. From the obtained nostril position, we estimated the region around the mouth and constructed a system that can recognize the opening and closing state of the mouth by binarization processing. As a result of the verification experiment by the subject, it confirmed that it is possible to detect the face area with "upward facing state", "sideways facing state", "left / right tilted state" which was impossible with the conventional recognition system, furthermore it showed that nostrils and mouth could be detected.

Chapter 4 describes the application of the detection system of facial feature points in the three-dimensional space constructed in the previous chapter to the control of the meal support manipulator. A spoon was provided at the tip of the three-link manipulator, and the system of the stereo camera developed in the previous chapter was put behind the three-link manipulator. The mouth coordinates of the three-dimensional space detected by the stereo camera was taken as the target position. The opening / closing state of the mouth was detected and used as a trigger to move the spoon attached to the manipulator close to the mouth. When the mouth opened state was recognized, the spoon approached to the mouth of the user, then temporarily stopped in front of the mouth. If the mouth kept opening state, the spoon entered the oral cavity. After the system confirmed that the mouth was closed, the spoon returned to its original position and a series of movements could be done.

Improvement of this system for that purpose further improves the measurement accuracy of the system, improvement of the detection algorithm which is not influenced by the brightness of the surrounding environment in the detection system, improvement of the detection rate, speeding up the processing speed. For manipulators as well, it is necessary to further increase the degree of freedom, to cope with movement in the Y-axial direction, and to expand the movable range.

References

[1] http://www8.cao.go.jp/shougai/whitepaper/h23hakusho/zenbun/pdf/h1/2_01.pdf

[2] http://www8.cao.go.jp/kourei/whitepaper/w-2010/gaiyou/22indexg.html

[3] J.Hammel, K.Hall, D.Lees, L. Leifer, M.V.Loos, I.Perkash, R.Crigler, 26-3, Journal of Rehabilitation Research and Development, Vol.26, No.3, pp.1-16,1989.

[4] J.R.Bach, Z Arie, W Charles, Wheelchair-Mounted Robot Manipulators-Long Term Use by Patients with Duchenne Muscular Dystrophy American Journal of Physical Medicine & Rehabilitation, Vol.69,No.2,pp.55-59,1990.

[5] S.Ishii, Meal-assistance Robot "My Spoon", Journal of Robotics Society of Japan, Vol.21, No.4, pp.378-381,2003.

[6] http://meetobi.com

[7]M. Topping and J. Smith, The Development of Handy1, A Robotic System to Assist the Severely Disabled. SixthInternational Conference on Rehabilitation Robotics 1999; 244-249.

[8] M. Topping, Handy1, A Robotic Aid to Independence for Severely Disabled People. Integration of Assistive Technology

in the Information Age 2001; 142-147.

[9] M. Topping, An Overview of the Development of Handy1, A Rehabilitation Robot to Assist the Severely Disabled. Journal of Intelligent and Robotic Systems 2002; 34: 253-263.

[10] J. Sijs, F. Liefhebber and G. Romer, Combined Position & Force Control for a robotic manipulator. IEEE 10th International Conference on Rehabilitation Robotics 2007; 106-111.

[11] G.Bradski and A.Kaehler: Learning OpenCV: Computer Vision with the OpenCV Library, O'REILLY press 2008

[12] P. Viola and Michael Jones: Robust real-time face detection, International Journal of Computer Vision, Vol.57, pp.137–154, 2004

[13] P. Viola and M. J. Jones: Detecting Pedestrians Using Patterns of Motion and Appearance, International Journal of Computer Vision, Vol.63, No.2, pp.153–161, 2005

[14] Y, Freund and R, E. Schapire, "A decisiontheoretic generalization of on-line learning and an application to boosting", Journal of Computer and System Sciences, No. 1, Vol. 55, pp. 119-139,(1997).

[15] Dalal. N, Triggs. B, "Histograms of Oriented Gradients for Human Detection", IEEE CVPR, pp. 886-893 (2005).

[16] R. E. Schapire, Y. Singer, "Improved Boosting Algorithms Using Confidence-rated Predictions", Machine Learning, No. 37, pp.297-336, (1999).

[17] D. Comaniciu, P. Meer, "Mean Shift: A Robust Approach toward Feature Space Analysis", IEEE PAMI, vol. 24, No. 5, pp. 603-619, (2002)

[18] E.Hinton, Osindero, S.and Teh, Y.-W.: A fast learning algorithm for deep belief nets, Neural Computation, Vol.18, pp.1527-1544(2006).

[19] P. Felzenszwalb, R. Girshick, D. McAllester and D. Ramanan,”Object Detection with Discriminatively Trained Part Based Models”,IEEE Transactions on PAMI,vol.32,no.9,pp.1627-1645,(2010).

[20] Stan Z. Li, Long Zhu, ZhenQiu Zhang, Andrew Blake, HongJiang Zhang, and Harry Shum. Statistical Learning of Multi-View Face Detection. In Proceedings of the 7th European Conference on Computer Vision. Copenhagen, Denmark. May, 2002.

[21] Rainer Lienhart and Jochen Maydt. An Extended Set of Haarlike Features for Rapid Object Detection. IEEE ICIP 2002, Vol. 1, pp. 900-903, Sep. 2002.

[22] A. Mohan, C. Papageorgiou, T. Poggio. Example-based object detection in images by components. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 4, pp. 349-361, April 2001.

[23] A. Mohan, C. Papageorgiou, T. Poggio. Example-based object detection in images by components. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 4, pp. 349-361, April 2001.

[24] C. Papageorgiou, M. Oren, and T. Poggio. A general framework for Object Detection. In International Conference on Computer Vision, 1998.

[25] Paul Viola and Michael J. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. IEEE CVPR, 2001.

[26] H. Rowley, S. Baluja, and T. Kanade. Neural network-based face detection. In IEEE Patt. Anal. Mach. Intell, Vol. 20, pp.22-38, 1998.

[27]Blessing of Dimensionality: High-dimensional Feature and Its Efficient Compression for Face Verification.

[28]P. Belhumeur, D. Jacobs, D. Kriegman, and N. Kumar. Localizingparts of faces using a consensus of exemplars. In CVPR, pages 545–552. IEEE, 2011. 2

[29]X. Cao, Y. Wei, F. Wen, and J. Sun. Face alignment by explicit shape regression. In Computer Vision and Pattern Recognition, pages 2887 –2894, June 2012. 1, 2

[30]Renliang Weng, Jiwen Lu, Junlin Hu, Gao Yang and Yap-Peng Tan, Robust feature set matching for partial face recognition, pp601-608, 2013

[31] B.Peng, M.Kano, N.Nakazawa, F.Wang, Y.Fujii, T.Yamaguchi, and T.Matsui: Detection of Nostril Position Based on Facial Color Distribution, Proc. ICTSS 2017, May 2017.

[32]Rama Chellappa, Charles L. Wilson, and Ssdd Dirohey: "Human and machine recognition of face", A survey. Proceeding of Th-e IEEE, Vol. 83, No 5, pages 704-740, 1995.

[33] Brand, J., and Mason, J., “A Comparative Assessment of Three Approaches to Pixel level Human Skin- Detection”. In Proc. of the International Conference on Pattern Recognition, vol. 1, 1056–1059, 2000.

[34] Brand, J., Mason, S., Roach, M., Pawlewski, M.,”Enhancing face detection in colour images using a skin probability map". Int. Conf. on Intelligent Multimedia, Video and Speech Processing, pp. 344-347, 2001.

[35] Brown, D., Craw, I., and Lewthwaite, J.,” A SOM Based Approach to Skin Detection

with Application in Real Time Systems”. In Proc. Of the British MachineVision Conference, 2001.

[36] Chai, D. and Bouzerdoum, A., “A Bayesian Approach to Skin Color Classification in YCbCr Color Space”. In Proc. Of IEEE Region Ten Conference, vol. 2, 421- 4124, 1999.

[37] Cho, M., Jang, H., and Hong, S., “Adaptive Skin Color Filter”, Pattern Recognition, Vol. 34, pp. 1067-1073, 2001.

[38] A. Jain, J. Bharti and M. K. Gupta: Improvements in OpenCV’s Viola Jones Algorithm in Face Detection - Tilted Face Detection, International Journal of Signal and Image Processing, No.5, pp.21-28, 2014

[39] Masahito TAKAHASHI, Yoshihiro TAKAYAMA, Takeshi NAGAYASU, Kennji TERABAYASHI, Kazunori UMEDA. Mouth Motion Recognition Using Shape Features and Low-resolution Images of Mouth Region.

[40] Mutsumi Watanabe, Natsuko Nishi.Research of Daily Conversation Transmitting System Based On Mouth Part Pattern Recognition, IEEJ Tran, EIS, Vol.124, No.3, 2004.

[41] Hironori Kai, Daisuke Miyazaki, Ryo Furukawa, Masahito Aoyama, Shinsaku Hiura, Naoki Asada. Speech Detection from Extraction and Recognition of Lip area, Vol.2011-CVIM-177 No.13, 2011.

[42] Takeshi Saitoh, Ryosuke Konishi. Lip Reading Based on Trajectory Feature. IEICE2007, Vol.J90-D, No.4, pp.1105-1114, 2007

[43] http://kondo-robot.com/product-category/servomotor/krs

[44] http://kondo-robot.com/

[45]Y. Kosaka and A. Shimada: Motion Control for Articulated Robots Based on Accurate Modeling, the 8th International Workshop on Advanced Motion Control (AMC2004), IEEE IE Society, 849/853 (2004)

[46] R. Paul: Robot Manipulators, Mathematics, Programming, and Control, MIT Press , 1981

[47] Y. Kosaka, A. Shimada and P. Viboonchaiceep: Vibration Control for Articulated Robots without Feedback of Disturbance Estimates, IECON2003, 849/853, IEEE Industrial Electronics Society (2003)

[48] T. Chen and B. Francis, Optimal Sampled-Data Control Systems, Communication and Control Engineering Series, Springer-Verlag (1995)

[49] Motohiro Kano, Position Measurement of Facial Feature Point in the three-dimensional space, Master thesis of Gunma University, 2012

[50] https://www.logicool.co.jp/ja-jp/video/webcams

[51] Manaf A. Mahammed, Amera I. Melhum, Faris A. Kochery, Object Distance Measurement by Stereo VISION, IJSAIT, Vol.2, No.2, pp.05-08 (2013)

[52] Transistor Technology Editor: To make robot's eyes, CQ Publisher, pp. 48-49 (2006)

[53] https://www.dh.aist.go.jp/database/head/index.html

[54] Hiroki Muguruma, Development of non-contact type interface using human sight, Master thesis of Gunma University, 2010

Acknowledgement

There is no denying that it is a tremendous challenge for me to do research and write a research paper in English. However, thanks to numerous people's tireless efforts, this paper has eventually come to fruition.

My first acknowledgements should go to Associate Professor Nobuaki Nakazawa of Gunma University, who taught me much technology and gave me many help, and thanks to the chief review teacher Professor Wei Shugang of Gunma University, and deputy review teachers, Professor Wang Feng of Maebashi Institute of Technology, Professor Zhu Chi of Maebashi Institute of Technology, and Associate Professor Toshikazu Matsui, thanks to give me advise in research and the paper.

And my second acknowledgement should go to Motohiro Kano of Gunma Industrial Technology Center, and the fellow students in the same laboratory of Nakazawa, who also gave me advise and helps.

Finally, I give my sincere thanks to my dear friends and fellow students, who have accompanied me throughout the years at the university, sharing sadness and happiness with me and teaching me a lot. Our friendship is always a treasure and a driving force to me.