Paper

Study on Modeling Effect of the Human Face on Monitoring Images

Mitsuhiro FUKUDA, Shunsuke KOUKO, Takuji UOZUMI and Kohji KAWAKAMI
Iwasaki Electric Co., Ltd. 3-12-4 Shiba Minato-city Tokyo, 105-0014 JAPAN

Received Octoober 24 2002, Accepted March 25, 2003

ABSTRACT

Television images become to be used for various fields as security observation in addition to broadcasting. It will be useful to install a camera system for security in order to obtain "restraint of a crime" and "criminal information". In particular, it is demanded to obtain "face discrimination" as criminal information. However, on an outdoor situation in the nighttimes there are many difficulties to obtain the "face discrimination" images by CCD camera. It is important how to establish lighting conditions to obtain video-monitored images. The lighting conditions are prepared by diversely modelled visual targets based on Cuttle's modelling evaluation experiments as reference and evaluating the level of "face discrimination" on the monitor. We discussed descriptive method for modelling effect using the illuminance as parameter.

KEYWORDS: modeling, CCD camera, security

1. Introduction

With the spread of CCD cameras, today, video images are in regular use not only for telecasting but also for security monitoring and various other purposes. In the security field, particularly, while more or less satisfactory video images can be obtained from CCD cameras (Charge Coupled Device cameras) under the interior lighting conditions. In general interior lighting provides no marked directionality of light, and consequently, the uniformity ratio of illuminance is high. Images taken outdoors, especially during the night, are often unsatisfactory in terms of physiognomy, i.e., in the interest of discrimination of facial features. This is because facial images are harshly shaded due to night time outdoor lighting condition where luminance contrast is normally very clear and usually involves marked directionality of light.

Generally speaking, it is believed that whether we can distinguish a clear figure or not on a video image depends largely on the camera's sensitivity, exposure and the luminance distribution on the subject. In addition it is also noted that whether facial features can be distinguished or not has a lot to do with the shades produced on a video-recorded face by the directionality of light (illumination vector). Actually, however, the minimum illuminance is the only lighting condition normally indicated in the specifications of video cameras. Establishment of a method to permit quantitative evaluation of the effect of shades on the video-recorded face images is therefore expected.

On the other hand, the past researches on the discrimination of face or facial features include the face modeling evaluation by Cuttle[1] using the ratio of illumination vector to scalar illuminance; and the facial look in relation to illuminance level of the street and security lighting by Miyamae and Takeuchi[2]. P. Rombauts, et al.[3] defined the minimum value of the semi-cylindrical illuminance required for face discrimination in residential areas, examining, at the same time, modeling evaluation using the ratio of vertical illuminance to semi-cylindrical illuminance.

These reports have provided knowledge and information of vital importance for further study of researchers on face discrimination and can also be used for the evaluation of monitored images. Notwithstanding, however, they have not definitely shown the relationship between modeling evaluation and monitored images, in particular, that between modeling evaluation and "face discrimination." We think a proper description of this relationship, if possible, will contribute to the design of security lighting above all.

With the above in mind, we have decided to study lighting conditions for obtaining monitored images that will permit "face discrimination" following the steps below:

(1) Evaluate the level of "face discrimination" on the monitor by preparing diversely modeled visual targets based on Cuttle's modeling evaluation experiments as reference.
(2) Examine the descriptive method for the modeling on the basis of the above evaluation results.

2. Preparation of visual targets

2.1 Image producing condition and lighting installation

In order to prepare visual targets having different illumination vector azimuths and altitudes and different ratios of illumination vector to scalar illuminance, a lighting installation as shown in Fig. 1 was installed in a dark room which permits interreflection to be ignored.

The room approximated a two-meter cube (2.0m wide×2.0m long×2.0m high), with an incandescent reflector lamp (500 W, 5,900 K) set up inside to create a vector illumination component. In addition a cubically assembled set of three band radiation fluorescent lamps (40 W, 5,000 K) were installed to produce scalar illuminance values. The entire room was blacked out with a curtain. At its center, a face a figure patterned after a person was set as a visual target. An opening (0.7 m × 1.2 m) was provided in the front panel for image taking and luminance measurements, and another (1.4 m × 1.2 m) in the back panel. This back opening together with a blackout curtain set 3.5 metres behind the opening enabled us to eliminate and therefore disregard possible effects of fluctuations in illumination vector and scalar illuminance on the background luminance.

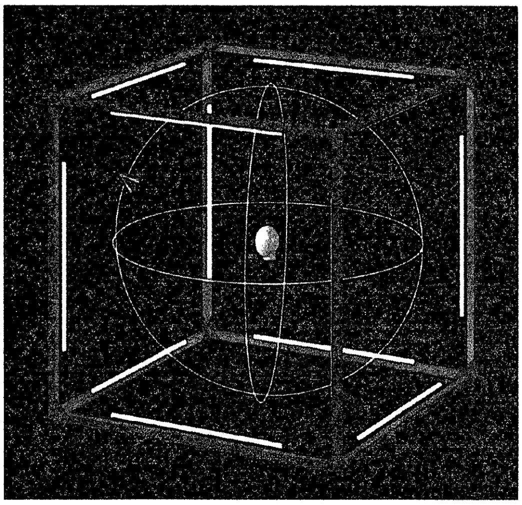

Fig.1 Lighting installation for obtaining visual target images.

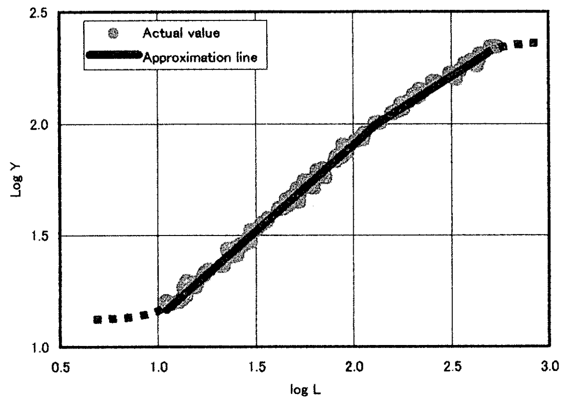

Fig.2 Correlation between luminance and gray-scale.

2.2 Conditions used for obtaining visual target images by digital still camera

A digital still camera (1.3 megapixels) having CCD elements built-in was used to obtain visual target images. Fig.2 shows the correlation between luminance and gray-scale for the camera which was used obtaining images. Lighting conditions were as follows: illumination vector was subjected to variables of 0, 30, 60 and 90 degrees in altitude (α) and to variables of 0, 30, 60, 90 and 120 degrees in azimuth (ϕ). At the same time, the incandescent reflector lamp used for the photograph and the lit fluorescent lamps were adjusted so that different ratios (0, 1.5, 2.5, 3.3 and 4.0) of illumination vector to scalar illuminance can be obtained.

The camera exposure was set at 90 degrees of illumination vector azimuth, a fixed ratio of illumination vector to scalar illuminance and also at a fixed angle of view of the face image (Fig. 3) so that only the difference in the shades caused by the variation in the illumination vector and scalar illuminance could be picked up.

Regarding the correct exposure, EV+1 and EV-1, these three types of images were obtained by changing only the illumination vector altitude and azimuth.

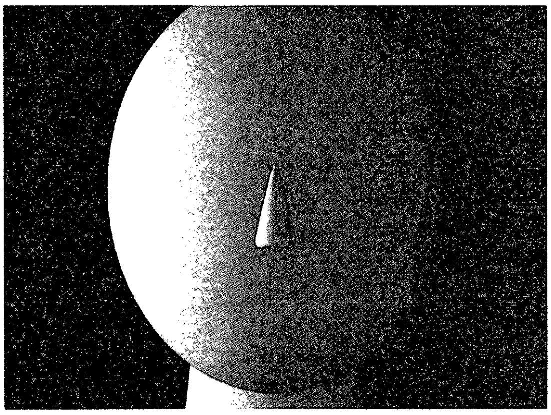

Fig.3 Size of a face for angle of view at exposure setting.

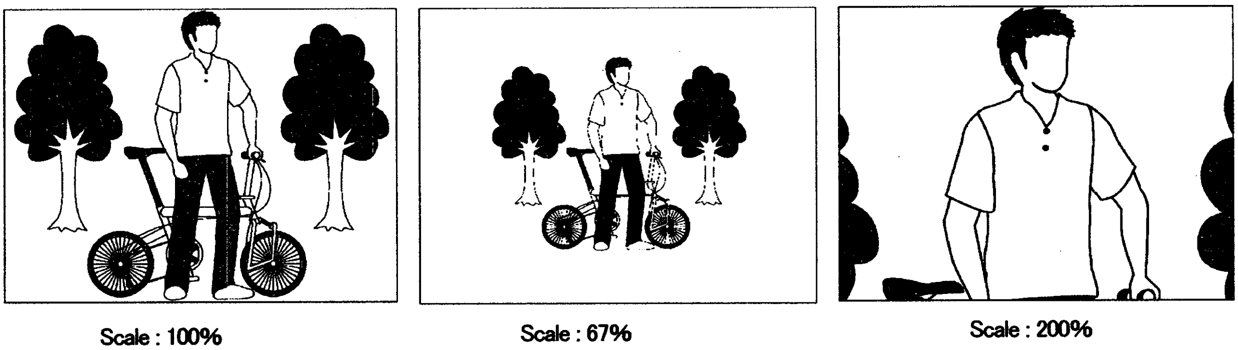

Fig.4 Reduced scale of a video image.

2.3 Size of visual targets

The discrimination of face or facial features also depends upon the angle of view of the face appearing on the image and the observation distance.

Reduced scale of a video-monitored person is generally expressed using the size of height as a person for angle of view, for example, wording such as "Scale is deemed to be 100% (Fig.4) where the monitor display and the person monitored have the same height". The most common reduced scale is between 50% and 67% for monitoring a scene of surroundings and human presence. And also 100% image is used discriminating the looks, appearance and manner of the monitored person, and used 200% images for discriminating physiognomy of the monitored person.

As stated above, estimates are that video monitors are normally working on a reduction scale of 50% to 67%. Accordingly, we set the size of the face appearing in the visual targets to be presented to experiment participants, to a reduced scale equivalent to 67%.

Table 1 Evaluation categories.

	The level of discrimination of a face
1	Shades are too harsh, the outline on human faces and facial features are almost impossible to recognize
2	Shades are harsh, the outline on human faces and facial features are barely recognizable
3	Shades are preferred, the outline on human face and facial features can be recognized clearly
4	Shades are soft, the outline on human faces and facial features are barely recognizable
5	Shades are too soft, the outline on human faces and facial features are almost impossible to recognize

3. Evaluation experiment for discrimination

3.1 Evaluation categories and the instructions given

Five evaluation categories (Table 1) were established according to the level of face discrimination. The following instructions and Fig. 5 as well as the explanation about the difference between “Softness and Harshness of shades on human faces” and “Bright and Dim images” were given to those who took part in the discrimination experiment:

<Instruction>

> Suppose that you are requested to discriminate the facial features of the criminal suspect monitored by a security camera. Various images of a person's face are coming on the monitor screens, so evaluate the level of discrimination of the individual appearing on the screens based on the following five categories.

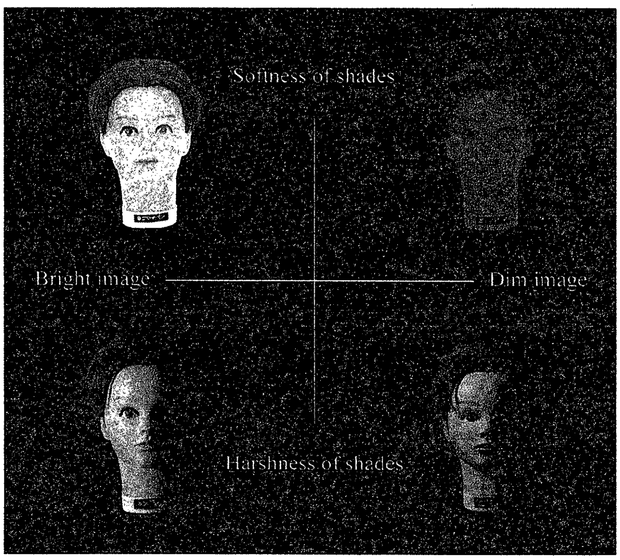

Fig.5 The difference between “Softness and Harshness of shades on human faces” and “Bright and Dim images”.

3.2 Evaluation condition

The visual targets were presented on two types of monitors: 17-inch CRT monitor (MITSUBISHI Diamondtron RD17GII) and 17-inch TFT liquid crystal color monitor (PROTON RD170C). As to the surroundings condition, the value of illuminance was set at 500lx on the desktop, and 250lx on the monitor screen. An achromatic color was used for the inside walls of the room so that the effect of colored reflected light from the surroundings was negligible and the light source is not reflected to the screens. Five measurements of the luminance at the center of each monitor screen were conducted at 1 degree of visual angle and from 1 meter of distance while the visual targets were shown on the screen, under the conditions of all white display and all black display. The average brightness was 92 cd/m^2 and 3.3 cd/m^2, respectively, for the CRT monitor, and 115 cd/m^2 and 2.8 cd/m^2, respectively, for the liquid crystal monitor.

3.3 Presentation of visual targets

A total of twenty-seven subjects having normal visual acuity after correction (Table 2) was grouped into six, and the same visual targets were presented on the CRT monitor and TFT liquid crystal monitor once each. After deciding the order of presentation using a table of random numbers, first, 8 cuts were shown to give the observers exercise, followed by the showing of a total of 124 cuts at intervals of 5 seconds of presentation and 5 seconds of pause. The observers had to give evaluation during each pause of 5 seconds.

Table 2 Configuration of subjects.

Sex/Age	20's	30's	40's	50's	Total
Male	5	10	3	4	22
Female	2	3	0	0	5
Total	7	13	3	4	27

Table 3 Distance between evaluation categories.

Interval scale	The level of discrimination of a face
1.00	Shades are too harsh, the outline on human faces and facial features are almost impossible to recognize
2.34	Shades are harsh, the outline on human faces and facial features are barely recognizable
3.00	Shades are preferred, the outline on human face and facial features can be recognized clearly
4.81	Shades are soft, the outline on human faces and facial features are barely recognizable
8.61	Shades are too soft, the outline on human faces and facial features are almost impossible to recognize

4. Results

The experimental results showed a normal distribution of almost all, the data obtained from 27 subjects for each of 124 cuts presented on CRT and TFT liquid crystal monitors, except those obtained from 2 subjects. For reporting the detailed results in this section, results from these 2 subjects were once each. After deciding the order of presentation using a table of random numbers, first, 8 cuts were shown to give the observers exercise, followed by the showing of a total of 124 cuts at intervals of 5 seconds of presentation and 5 seconds of pause. The observers had to give evaluation during each pause of 5 seconds.
excluded from the data reduction.

4.1 Category-to-category distance

The evaluation categories (Table 1) were converted into interval scales by the method of successive categories[5)6)]. Table 3 shows the interval scales where scale 1.00 represents an image in which " Shades are too harsh, the outline on human faces and facial features are almost impossible to recognize," and scale 3.00, the one in which " Shades are preferred, the outline on human face and facial features can be recognized clearly."

4.2 Difference in the evaluation results due to the type of monitor

To verify possible difference in the evaluation results according to the type of monitor used (CRT and liquid crystal monitors), the mean value was calculated for each visual target based on the interval scale obtained by the method of successive categories as described above. Fig. 6 shows the mean value of each visual target. In the chart,

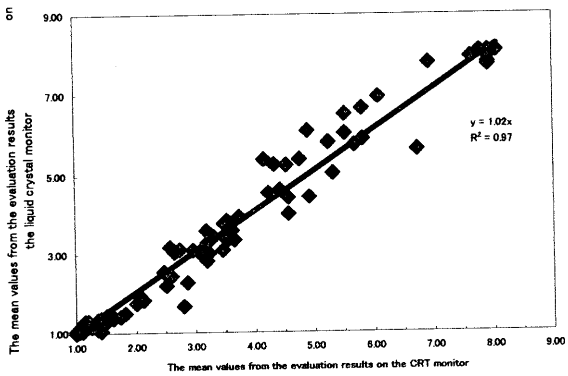

Fig.6 The mean value of each evaluation results on the CRT monitor and the liquid crystal monitor.

the mean values from the evaluation results on the CRT monitor are plotted in cross axis, and those from the evaluation results on the liquid crystal monitor in vertical axis. The resultant determinant coefficient of correlation was 0.97. Then, we conducted the significance test whether or not there is a significant difference between the evaluation results from the two monitors using the population mean difference. No significant difference was shown.

As it was verified that there was no significant difference between the evaluation results obtained using the CRT monitor and those obtained using the liquid crystal monitor, we decided to bring them together for the purpose of the analysis described below.

4.3 Representation by vector altitude and ratio of illumination vector to scalar illuminance

As the face modeling evaluation had already been researched by Cuttle[1] using the ratio of illumination vector to scalar illuminance, we plotted the evaluation results of the 124 stimulant cuts after the model proposed by Cuttle. In Fig. 7 through to Fig. 21, the illumination vector altitude is plotted in cross axis, and the ratio of illumination vector to scalar illuminance in vertical axis, and the evaluated images and their mean values are plotted using the illumination vector azimuth and the camera exposure as parameters.

5. Analysis

5.1 Comparison with Cuttle's research

The use of the illumination vector azimuth (ϕ) and altitude (α) and the ratio of illumination vector to scalar illuminance in the same way as with Cuttle's research led to a distribution of two different plot groups: one is composed of visual targets which permit face discrimination, and the other, of those which do not (Fig. 7 to Fig. 21). This suggests that the modeling effect of the monitored face image can also be described using these three parameters.

However, whereas Cuttle's research provides the evaluation result of "Shades too harsh" when the value of the ratio of illumination vector to scalar illuminance is large, our research data show some evaluation result of "Shades too soft" when the illumination vector azimuth is at 120 degrees. These refer to the images picked up against the light and the said evaluation result of "Shades too soft" is considered to be attributable to the fact of no shading around eyes, nose and mouth due to the effect of such backlight. This phenomena seems to be caused by the exposure at the time of image capture and because our visual targets are monitored images.

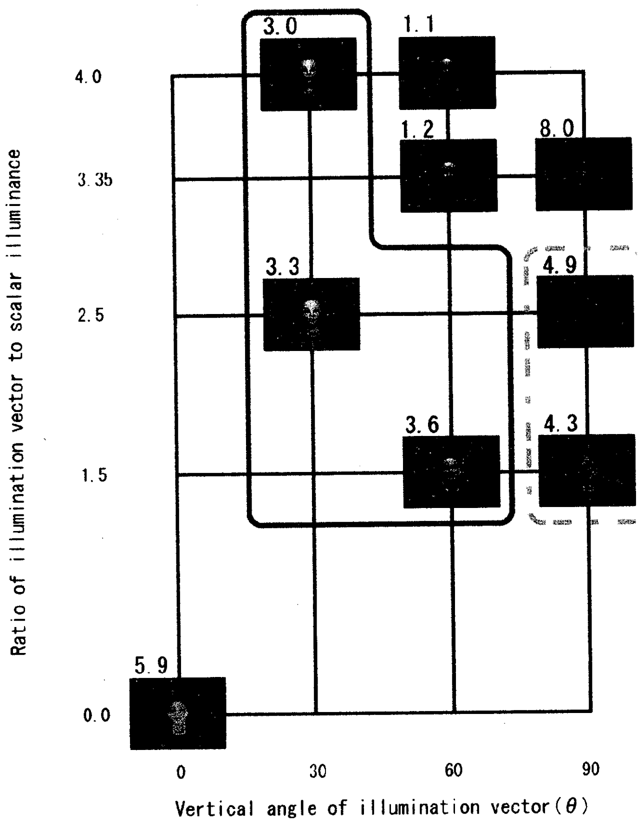

Fig.7 Horizontal angle : 0 degree (Exposure : Appropriate).

NOTE:

The area of Blue continuous line:

Visual targets, which evaluated "Shades are preferred, the outline on human face and facial features can be recognized clearly"

The area of Green dotted line:

Visual targets, which evaluated "Shades are harsh, the outline on human faces and facial features are barely recognizable"

The area of Orange alternate long and two short dashes line:

Visual targets, which evaluated "Shades are soft, the outline on human faces and facial features are barely recognizable"

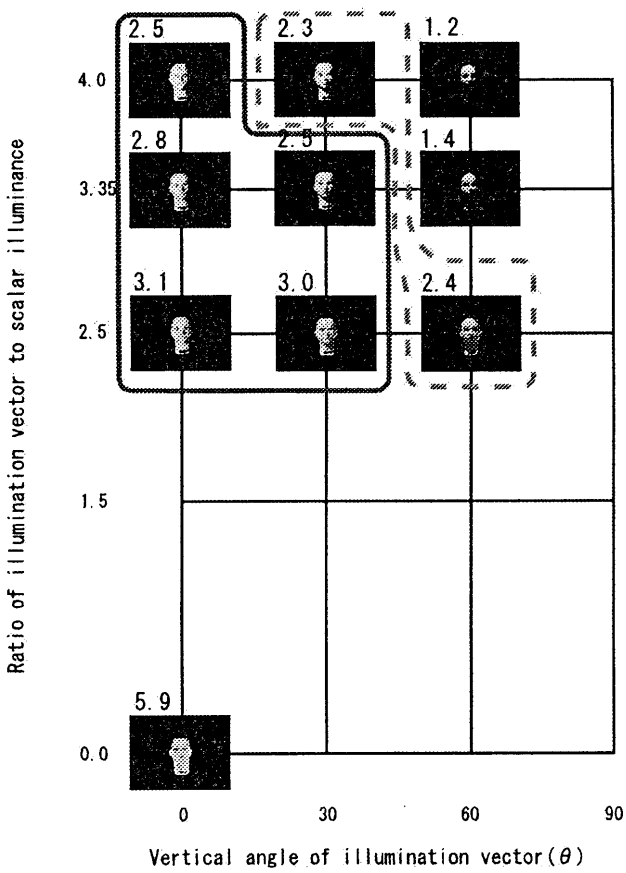

Fig.8 Horizontal angle : 30 degree (Exposure : Appropriate).

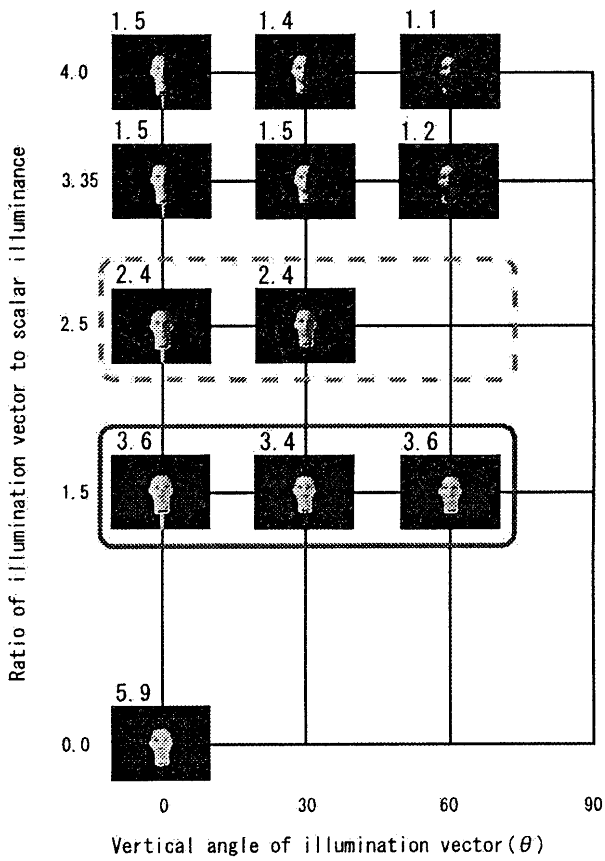

Fig.9 Horizontal angle : 60 degree (Exposure : Appropriate).

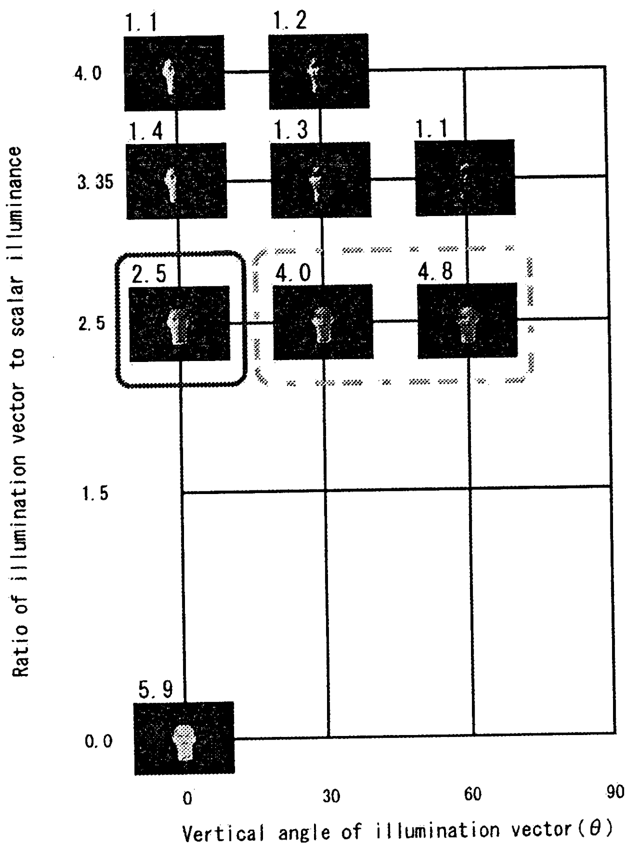

Fig.10 Horizontal angle : 90 degree (Exposure : Appropriate).

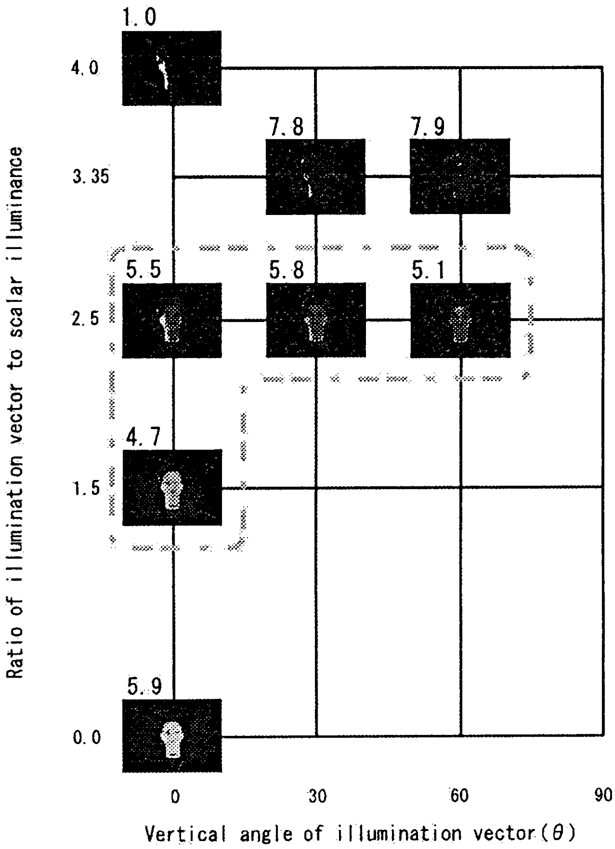

Fig.11 Horizontal angle : 120 degree (Exposure : Appropriate).

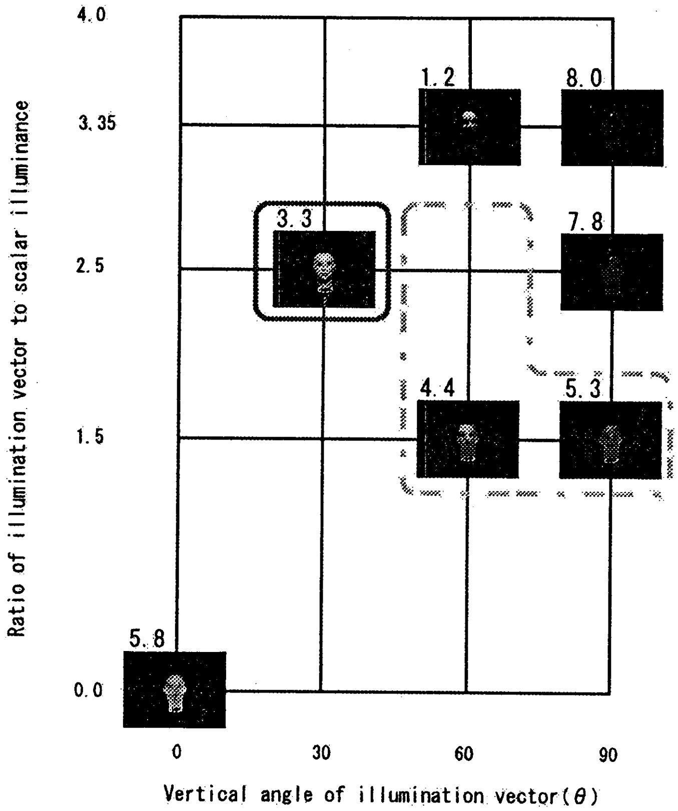

Fig.12 Horizontal angle : 0 degree(Exposure : EV−1).

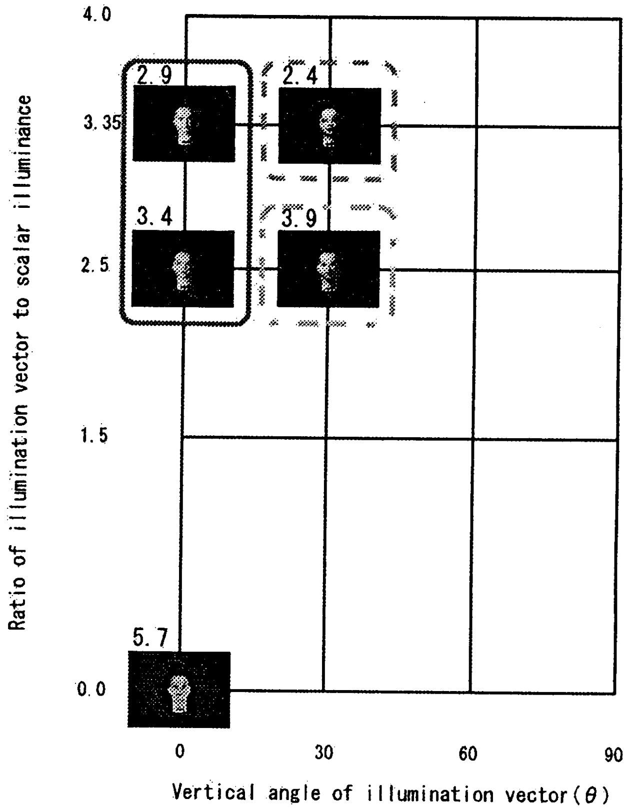

Fig.13 Horizontal angle : 30 degree(Exposure EV−1).

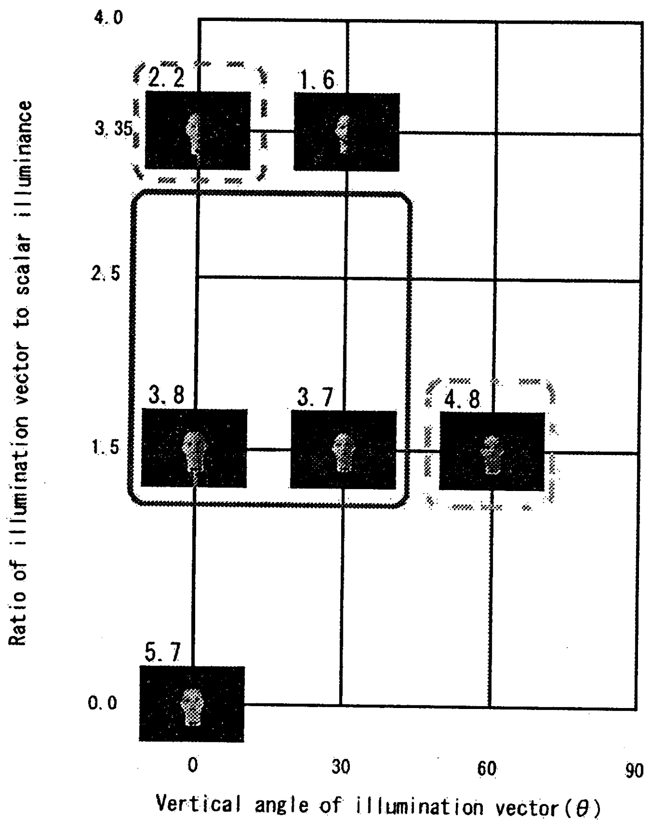

Fig.14 Horizontal angle : 60 degree(Exposure : EV−1).

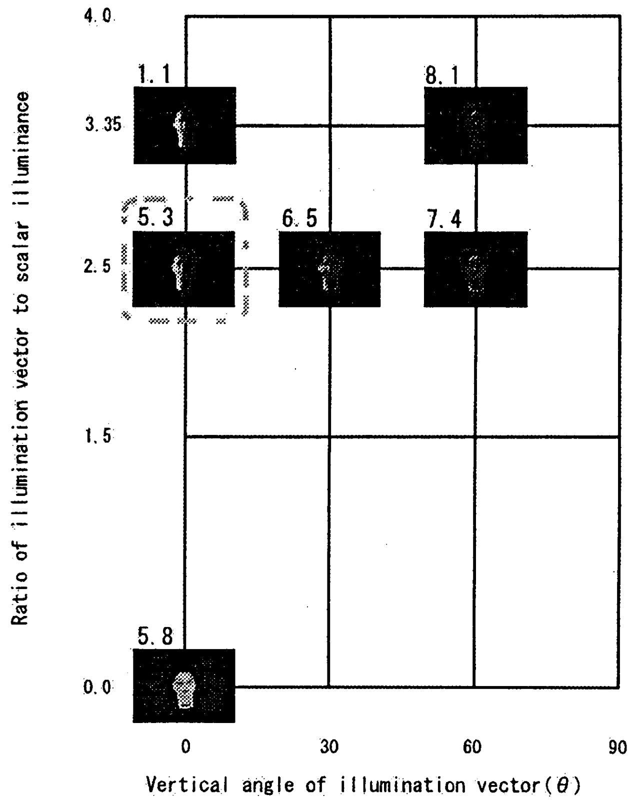

Fig.15 Horizontal angle : 90 degree(Exposure : EV−1).

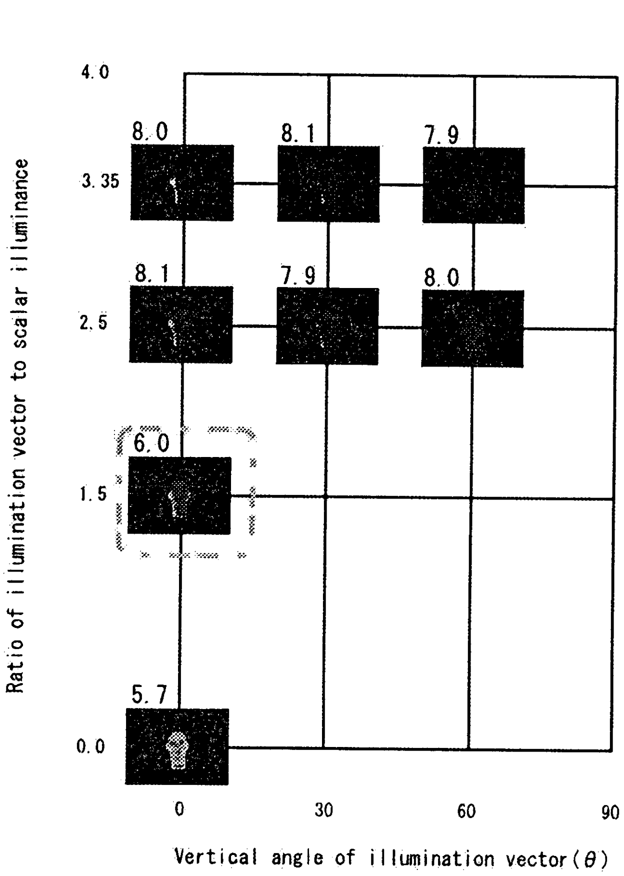

Fig.16 Horizontal angle : 120 degree (Exposure : EV−1).

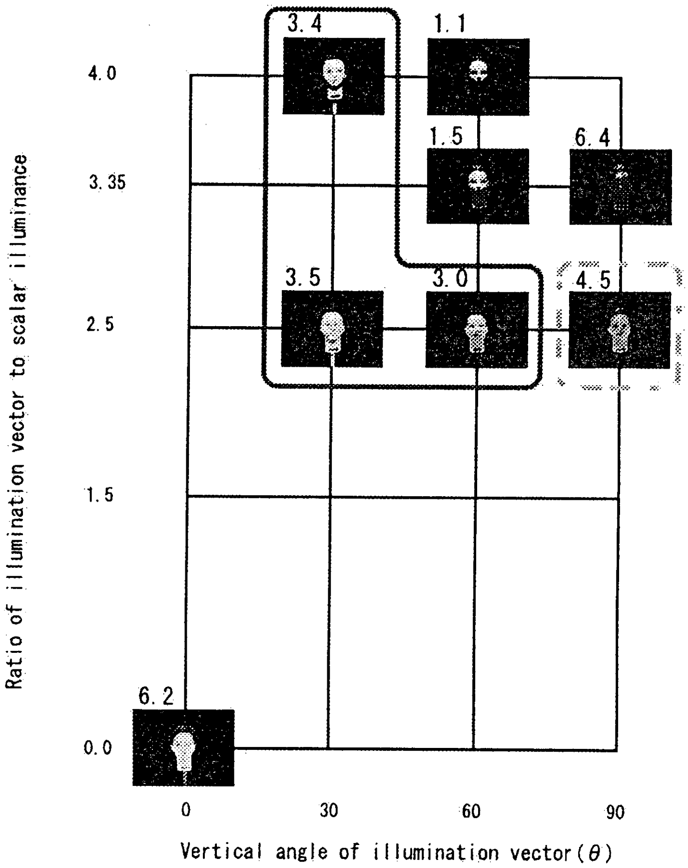

Fig.17 Horizontal angle : 0 degree (Exposure : EV+1).

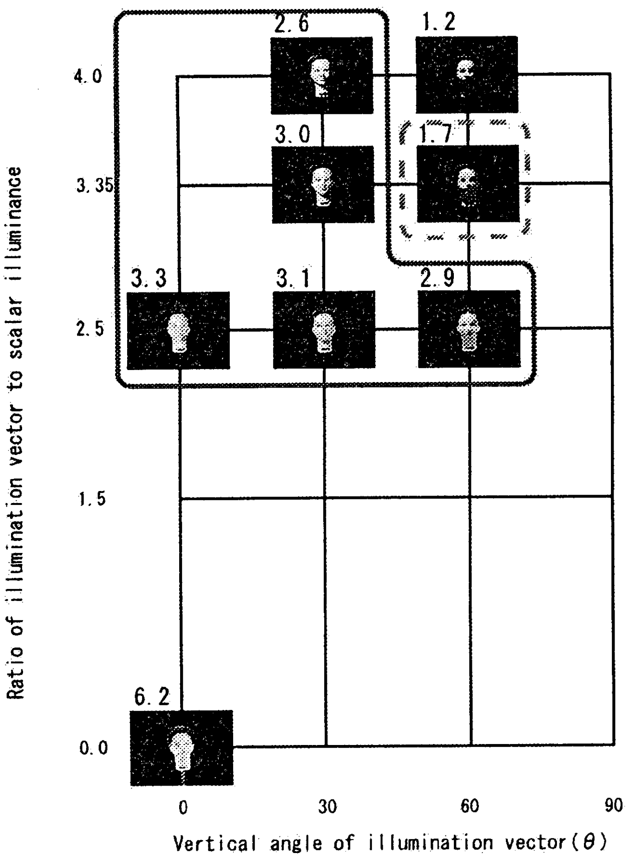

Fig.18 Horizontal angle : 30 degree (Exposure : EV+1).

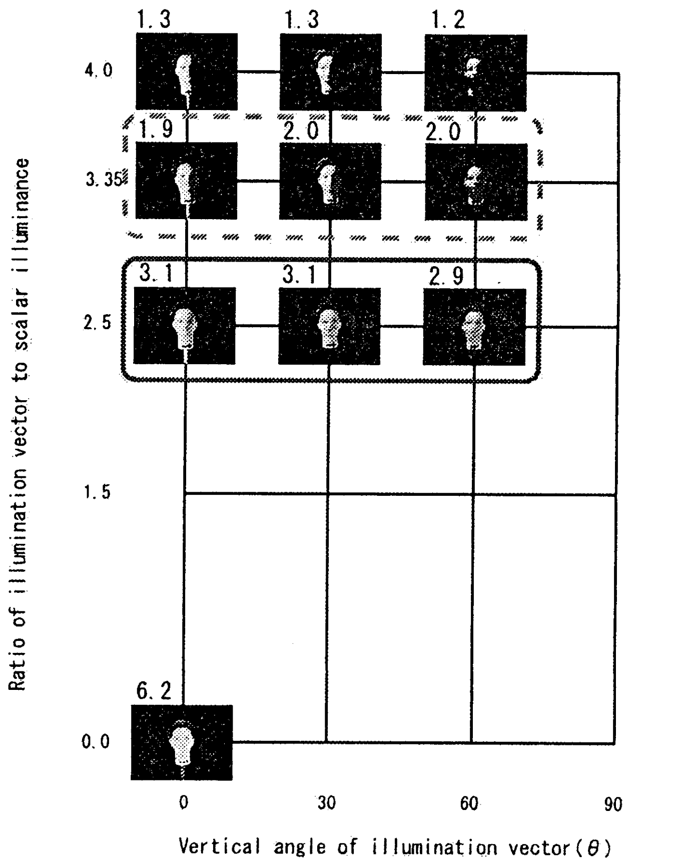

Fig.19 Horizontal angle : 60 degree (Exposure : EV+1).

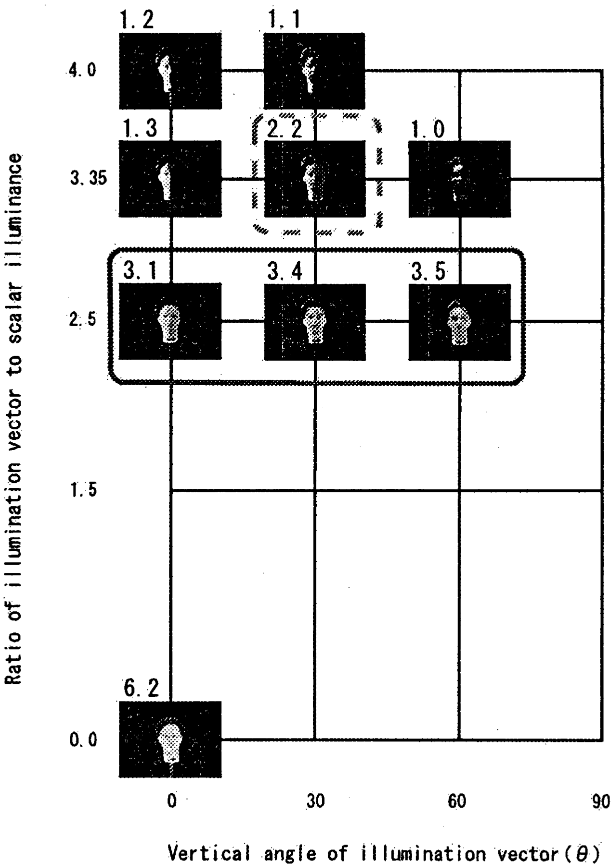

Fig.20 Horizontal angle : 90 degree (Exposure : EV+1).

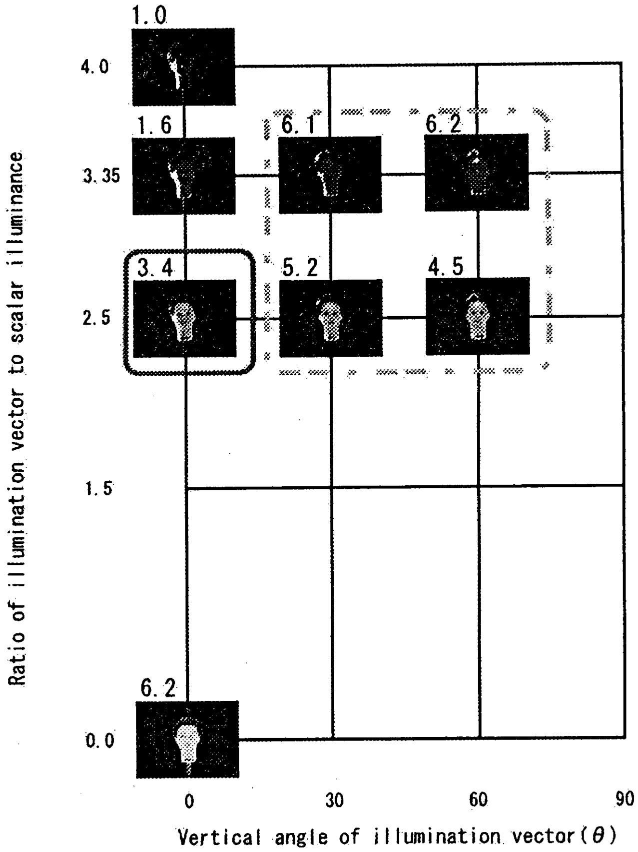

Fig.21 Horizontal angle : 120 degree (Exposure : EV+1).

5.2 Difference of exposure level at the time of image capture

The relation between exposure levels (underexposure: EV-1, appropriate exposure, and overexposure: EV+1) and evaluation results is described below.

Arranged by horizontal angle in Fig. 22, Fig. 23, Fig. 24 and Fig. 25 are the changes in evaluation score caused by different exposure levels to the visual targets which were taken using the same vectors (azimuth (ϕ) and altitude (α)) and the same ratio of illumination vector to scalar illuminance.

There are two marked groups: visual targets whose evaluation was not affected practically by the change of exposure level (EV-1, appropriate, EV+1) (Fig. 22 and Fig. 23), and those whose evaluation was seriously affected (Fig. 24 and Fig. 25), demonstrating a difference between the two evaluation scores.

Analyzing their respective relations with visual targets, those in the second group were evaluated with "Shades too soft" at EV-1 level because the face image was dim. As the exposure changed to the appropriate and overexposure (EV+1) levels, the face image got gradually brighter, resulting in better evaluation; that is, the brightness of the face was the principal element of evaluation for this group. On the other hand, importance was seemingly placed on the shades in the evaluation of those in the first group. In short, it can be said that the evaluation based on the shades is not affected practically by the exposure level.

5.3 Descriptive method for modeling effect

Very complicated works are involved in a lighting design when we try to obtain good monitored images using the above-mentioned three parameters, namely, illumination vector azimuth (ϕ), altitude (α) and the ratio of illumination vector to scalar illuminance. It is desired to provide a clear description of lighting conditions using the value of illuminance as parameter as it can be easily handled in lighting design.

With this in view, we attempted to describe

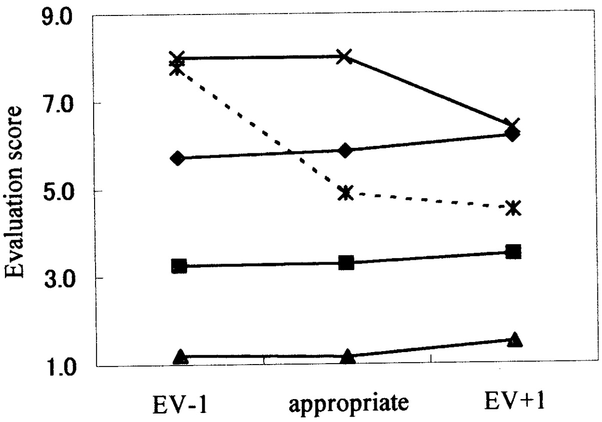

Fig.22 Relation between exposure levels and evaluation results (Horizontal angle : 0 degree).

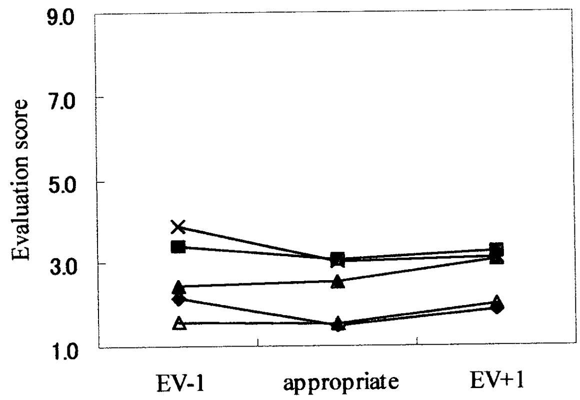

Fig.23 Relation between exposure levels and evaluation results (Horizontal angle : 30 and 60 degree).

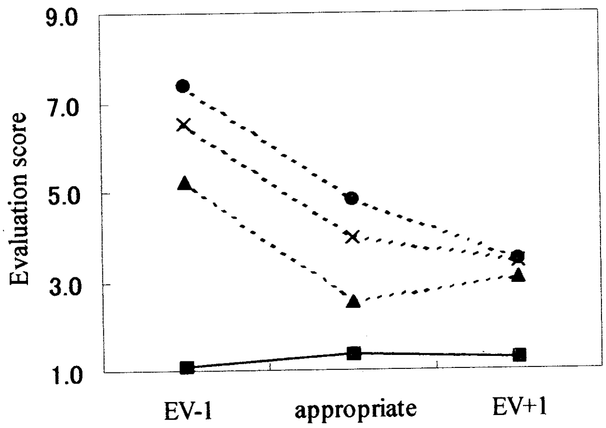

Fig.24 Relation between exposure levels and evaluation results (Horizontal angle : 90 degree).

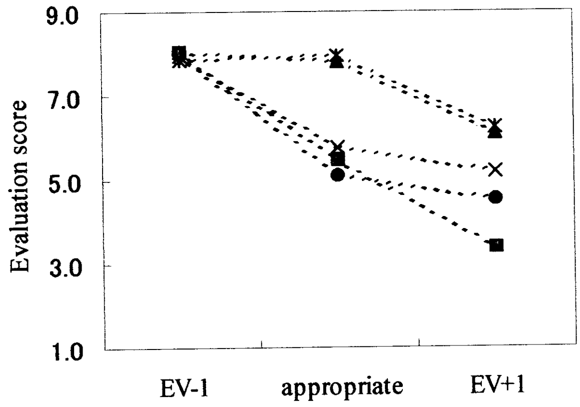

Fig.25 Relation between exposure levels and evaluation results (Horizontal angle : 120 degree).

modeling effect on the basis of the six-faces-of-cube illuminance, a concept developed by F. Haeger. The illumination vector is divided into horizontal and vertical components based on its azimuth (ϕ) and altitude (α). Components are disintegrated into two directions by taking the difference between horizontal illuminance in the upward/downward directions as the value of "Illuminance of a vertical direction", and the difference between vertical illuminance on the right and left sides of the face as the value of "Illuminance of a horizontal direction". The mean spherical illuminance was selected as the scalar illuminance of the light incident upon the image head total luminous flux while the semi-cylindrical illuminance toward the camera was selected as the quantity contributing to face discrimination. Supposing that directions of light incident upon six planes represent the azimuths of illuminance, these illuminance can be formulated by the following expressions[7]:

$$E_s = (E_{hu} + E_{hd} + E_{vf} + E_{vr} + E_{vb} + E_{vl})/6 \quad \text{..........(1)}$$

$$E_{sc} = (E_{vf} + E_{vr} + E_{vb} + E_{vl})/4 + E_{vf} + E_{vb})/\pi \quad \text{..........(2)}$$

$$(E_{vf} - E_{vb}) = \vec{E} \cdot \cos\alpha \cdot \cos\phi \quad \text{..........(3)}$$

$$(E_{vf} - E_{vl}) = \vec{E} \cdot \cos\alpha \cdot \sin\phi \quad \text{..........(4)}$$

$$(E_{hd} - E_{hu}) = \vec{E} \cdot \sin\alpha \quad \text{..........(5)}$$

where:

E_s: Mean spherical illuminance

E_{sc}: Semi-cylindrical illuminance

$\vec{E}$: Illumination vector

E_{hu}, E_{hd}: Horizontal illuminance in the upward/downward directions

E_{vf}, E_{vr}, E_{vb}, E_{vl}: Vertical illuminance in four directions (front, back, right and left)

Examinations were made in an effort to obtain a descriptive method for modeling effect by converting lighting conditions at the time of obtaining visual targets images into horizontal illuminance in the upward/downward directions (E_{hu}, E_{hd}), vertical illuminance in four directions (E_{vf}, E_{vr}, E_{vb}, E_{vl}), mean spherical illuminance (E_s) and semi-cylindrical illuminance (E_{sc}) using the above expressions.

As a result, we could obtain a scatter diagram that permits a clear distinction between the visual targets allowing face discrimination and those not,

Table 4 The examination result of the description method for the modeling effect

	$(E_{vf}-E_{vb})/E_s$	$(E_{vr}-E_{vl})/E_s$	$(E_{vf}-E_{vb})/E_{sc}$	$(E_{vr}-E_{vl})/E_{sc}$
$E_{hd}-E_{hu}$	×	×	×	×
E_{hd}/E_s	×	×	×	×
E_{hd}/E_{sc}	×	×	×	×
$(E_{hd}-E_{hu})/E_s$	×	O	×	O
$(E_{hd}-E_{hu})/E_{sc}$	×	×	×	×

Table 5 Distance between evaluation categories.

Classification	The level of discrimination of a face	Interval scale
1	Shades are preferred, the outline on human face and facial features can be recognized clearly	2.5～3.8
2	Shades are harsh (Soft), the outline on human faces and facial features are barely recognizable	6～2.5, 3.8～6.3
3	Shades are too harsh (Soft), the outline on human faces and facial features are almost impossible to recognize	0～1.6, 6.3～8.6

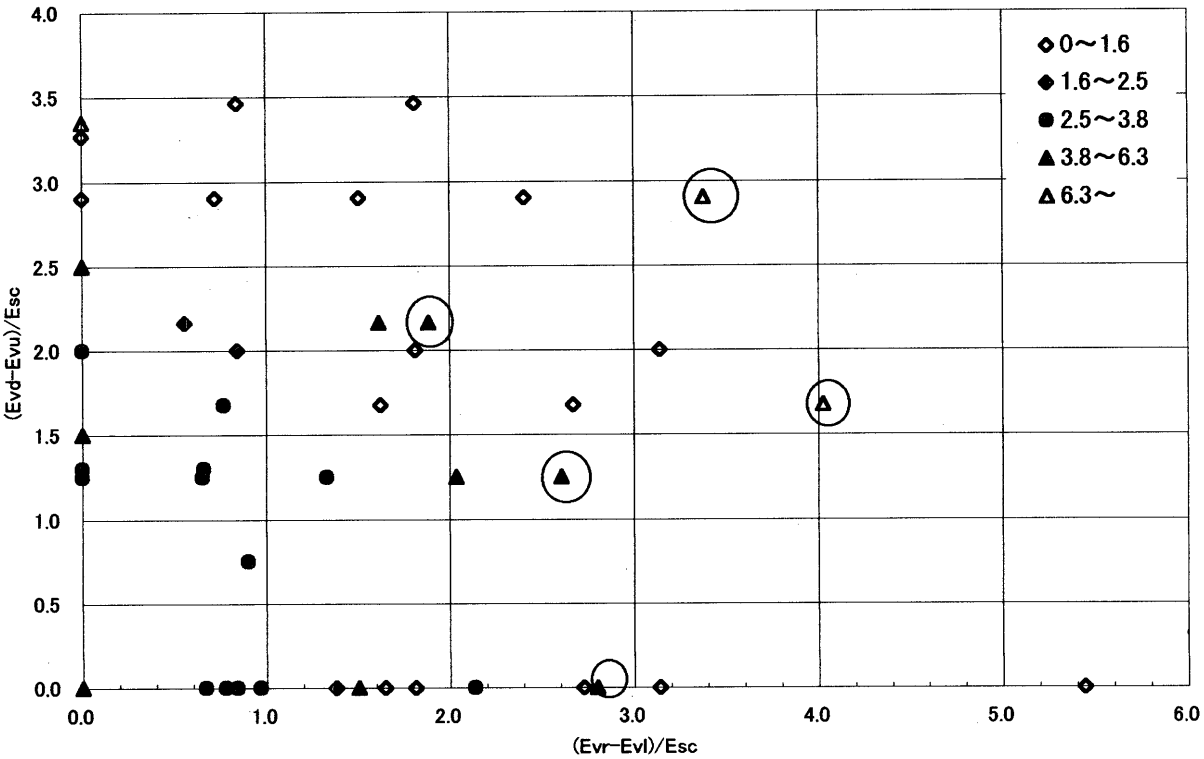

Fig.26 Plotted figure of evaluation result.

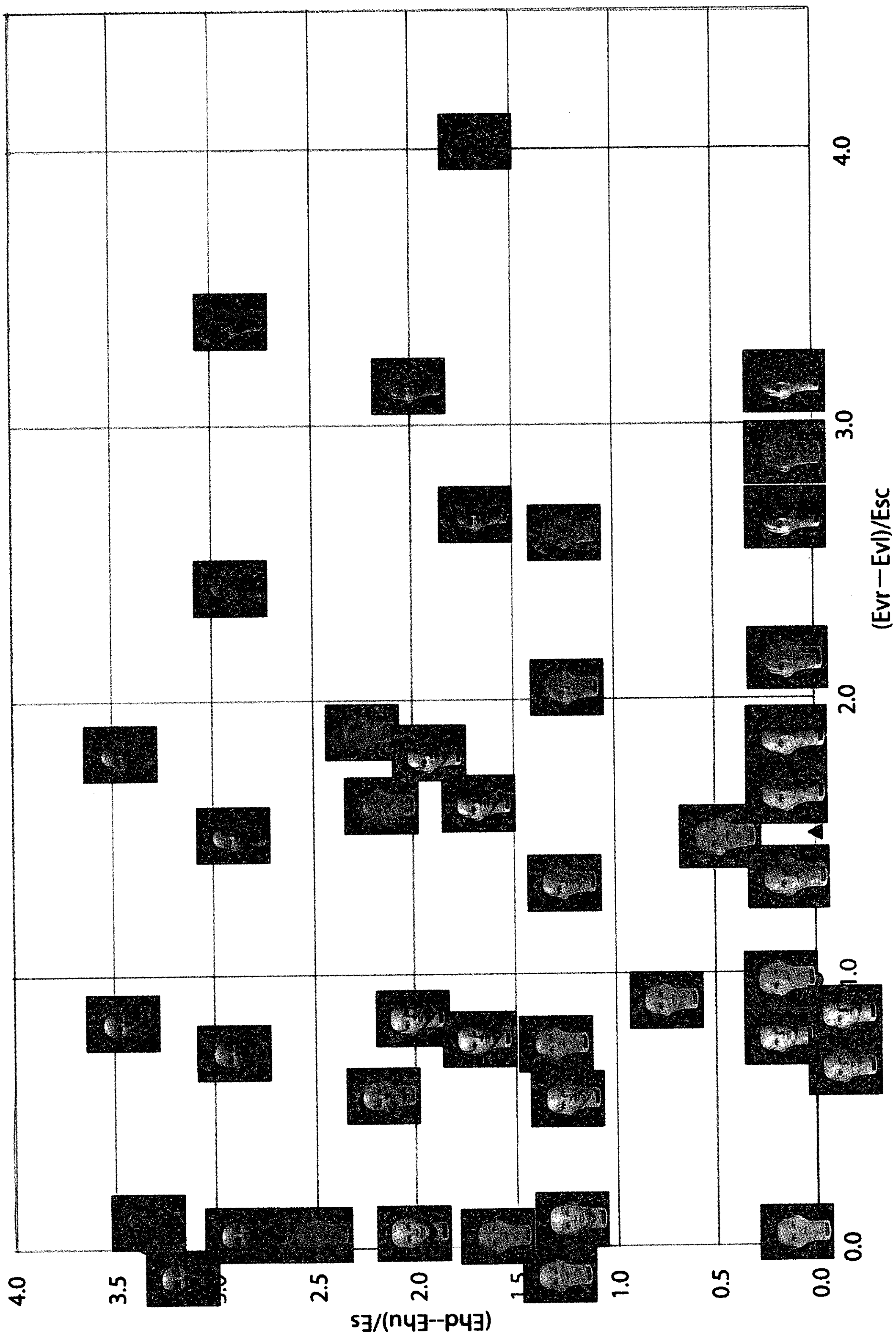

Fig.27 Plotted figure of evaluation result (Image)

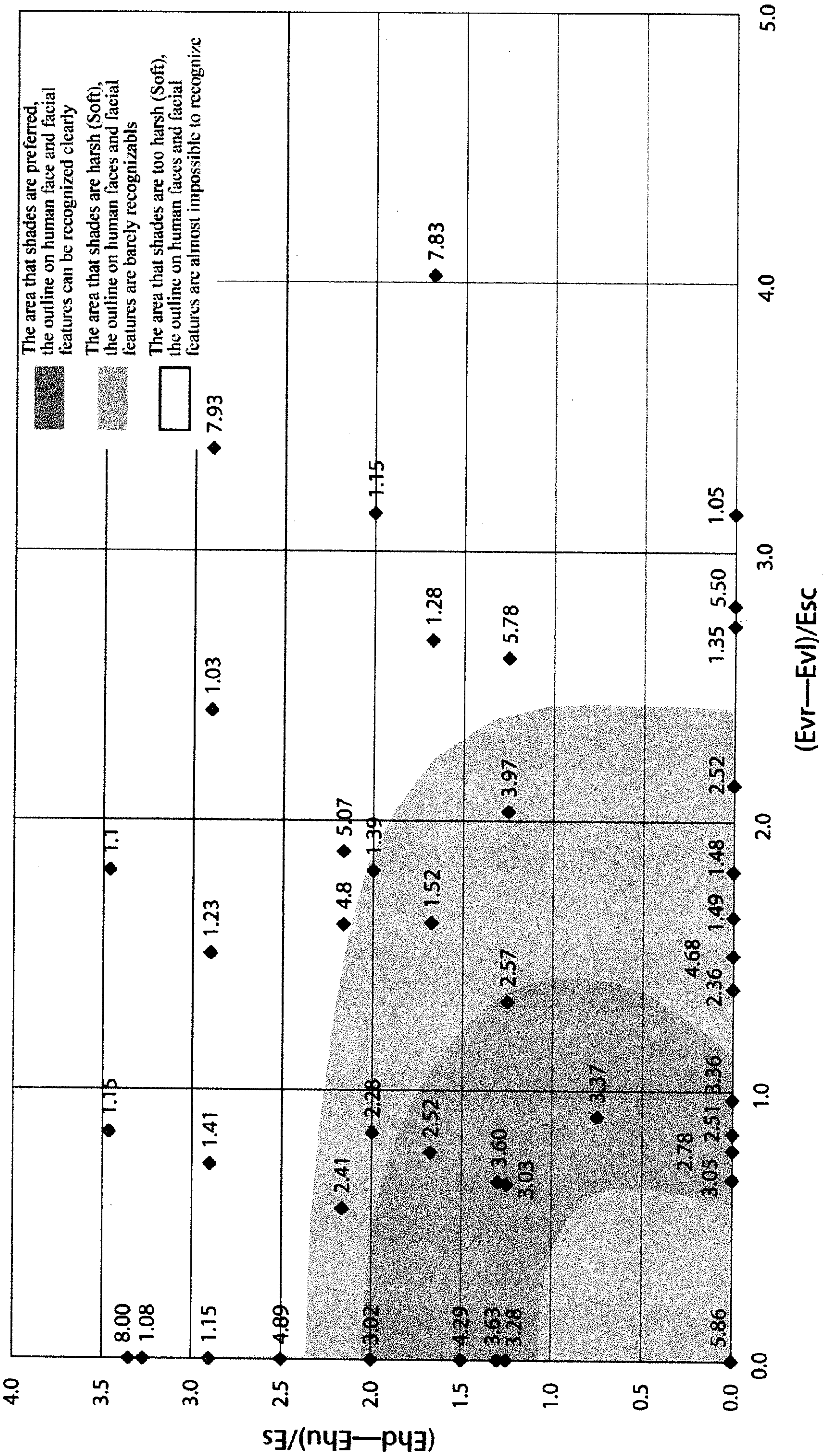

Fig.28 The relation between monitoring image and modeling effect of human face.

by combining expressions "$(E_{hd}$ - $E_{hu})/E_s$" and "$(E_{vr}$ - $E_{vl})/$ E_{sc} " or "$(E_{vr}$ - $E_{vl})/E_s$" as shown in Table 4. When the mean spherical illuminance expressed by "$(E_{vr}$ - $E_{vl})/E_s$" is used, however, there is no distinction between 120- and 60-degree azimuths if the size of illumination vector is the same. This suggests that the use of the semi-cylindrical illuminance is preferable.

Fig. 26 and Fig. 27 show the scatter diagrams of visual targets plotted using "the ratio of mean spherical illuminance to the difference between horizontal illuminance in the upward/downward directions $(E_{hd}$ - $E_{hu})/Es$" and "the ratio of semi-cylindrical illuminance to the difference between vertical illuminance on the right and left sides of the face $(E_{vr}$ - $E_{vl})/E_{sc}$."

In Fig. 26, the evaluation results were classified into five categories in accordance with the evaluation scores indicated in Table 3, showing a clear distinction between the visual targets allowing face discrimination and those not. The visual targets enclosed with solid line circles are those with 120-degree illumination vector azimuth which were judged with "Shading too soft" as they referred to the images picked up against the light and there was no shading around eyes, nose and mouth due to the effect of such backlight. These dim face images also came in the perimeter of the chart.

On the other hand, Fig. 27 plots visual targets as are. It is verified that a satisfactory distribution can be obtained according to the "level of discrimination of the video-monitored person" in these visual targets when the proposed two parameters are used.

Shown in Fig. 28 are the levels of discrimination of the video-monitored person, classified by the categories specified in Table 5, based on the results from the above examinations.

6. Conclusion

The purpose of this study was to establish lighting conditions to obtain video-monitored images that will permit "face discrimination" by preparing diversely modeled visual targets based on Cuttle's modeling evaluation experiments as reference and evaluating the level of "face discrimination" on the monitor. Following on from these evaluation results, we discussed descriptive method for modeling effect using the illuminance as parameter.

As a result, we could bring the following to light and make it possible to estimate at the time of image capture whether or not satisfactory images could be obtained:

a) There was no significant difference between the evaluation results obtained using the CRT monitor and those obtained using the liquid crystal monitor.

b) The modeling effect of the video-monitored face can be described using "the ratio of mean spherical illuminance to the difference between horizontal illuminance in the upward/downward directions $(E_{hd}$ - $E_{hu})/E_s$" and "the ratio of semi-cylindrical illuminance to the difference between vertical illuminance on the right and left sides of the face $(E_{vr}$ - $E_{vl})/E_{sc}$."

c) The lighting conditions to obtain good images of the face or facial features are as shown in Fig. 28 to the extent of this study. The range of these conditions can be roughly summarized as below:

> The solution of $(E_{hd}$ - $E_{hu})/E_s$ is to fall within the range between 1.0 and 2.0 and that of $(E_{vr}$ - $E_{vl})/E_{sc}$ should be below 1.3, or
> the solution of $(E_{hd}$ - $E_{hu})/E_s$ is to be below 1.0 and that of $(E_{vr}$ - $E_{vl})/E_{sc}$ should fall within the range between 0.6 and 1.3.

References

(1) Cuttle, C.: Lighting Patterns and the Flow of Light, Light. Res. Technol., 3, pp.171-189(1971).
(2) Atsuko Miyamae, Tetsuji Takeuchi : Illuminance to be provided for residential roads, J.Illum.Engng.Inst.Jpn. 73-6, pp.25-29(1989).
(3) P Rombauts, H Vandewyngaerde, G Maggetto: Minimum semicylindrical illuminance and modeling in residential area lighting, Light. Res. Technol., 21-2, pp.49-55(1989).
(4) Shino Okuda, Ryuji Satoh : Study on Establishing an Evaluation Method for the Visibility of the Human Face. J.Illum.Engng.Inst.Jpn. 84-11, pp.25-29(2000)
(5) J.P. Guilford : Psychometric Methods, Baifuukan, pp.276-301(1959).
(6) Yoshihisa Tanaka : Shinrigakutekisokuteihou (measurement method by psychology) . University of Tokyo Press, pp.119-130(1968).
(7) F. Haeger, Berlin: Über die Messung von lichtfeldgrößen zur räumlichen Bewertung der Beleuchtung mit einfachem Meßinstrumentarium, LICHTTECHNIK, 26.Jahrgang Nr.9, pp.385-(1974).