

# Head Position and Pose Model and Method for Head Pose Angle Estimation based on Convolution Neural Network

Kohei Arai, Akifumi Yamashita, Hiroshi Okumura

Faculty of Science and Engineering, Saga University, Saga City, Japan

**Abstract**—Head position and pose model is created. Also, a method for head poses angle estimation based on Convolution Neural Network (CNN) is proposed. 3D head position model is created from these locations and obtain 3D coordinate of head position. The method proposed here uses CNN. As for the head pose detection, OpenCV and Dlib of the open-source software tools are used with Python program. The images used were RGB images, RGB images + thermography, grayscale images, and RGB images assuming images obtained by near infrared rays, with only the red channel elements extracted. As a result, the RGB image model was the most accurate, but considering the criteria set, the RGB image model was used for morning and daytime detection, and the near-infrared image was used for nighttime and rainy weather scenes. It turned out that it is better to use the model obtained by the training in. The experimental results show almost perfect head pose detection performance when the head pose angle ranges from 0 to 180 degrees with 45 degrees steps.

**Keywords**—CNN; head pose; OpenCV; Dlib; open-source software; python

## I. INTRODUCTION

Head movement detection has received significant attention in recent research. One of the specific purposes for head movement detection and tracking is to allow the user to interact with a computer or new devices like mobile phone. The increased popularity of the wide range of applications of which head movement detection is a part, such as assistive technology, virtual reality, and augmented reality, have increased the size of research aiming to provide robust and effective techniques of real-time head movement detection and tracking [1].

Most of the head pose estimation method is based on computer vision approach, like [2], [3]. Liu et al. [2] introduced a video-based technique for estimating the head pose and used it in an image processing application for a real-world problem; and attention recognition for drivers. Murphy-Chutorian and Trivedi presented a static head pose estimation algorithm and a visual 3-D tracking algorithm based on image processing and pattern recognition<sup>1</sup>. Kupetz et al. [3] implemented a head movement tracking system using an IR camera and IR LEDs.

Another approach for head movement detection is by using sensors such as gyroscopes and accelerometers. King et al. [4] implemented a hands-free head movement classification system which uses pattern recognition techniques with

mathematical solutions for enhancement. A dual axis accelerometer mounted inside a hat was used to collect head movement data. A similar method was presented by Nguyen et al. [5]. The method detects the movement of a user's head by analyzing data collected from a dual-axis accelerometer and pattern recognition techniques. But still no application based on the proposed method was suggested. Other sensor-based approaches are like [6], [7]. However, it needs more theoretical proofs and more experiments and accuracy analysis.

A combination of different techniques can be used in head tracking systems. Satoh et al. [8] proposed a head tracking method that uses a gyroscope mounted on a head mounted device (HMD) and a fixed bird's-eye view camera responsible for observing the HMD from a third person viewpoint.

In our previous research work, we propose head movement detection and tracking as a controller for 3D object scene view [9] and the combination of user's head and body movement as a controller for virtual reality labyrinth game [10]. One of the problems of the previous method for head pose angle estimation is week accuracy.

In this paper, 3D head position model is created from these locations and obtain 3D coordinate of head position. Then, the method proposed here uses Convolutional Neural Network (CNN) in order to improve head pose angle estimation accuracy. As for the head pose detection, OpenCV and Dlib<sup>2</sup> of the open-source software tools are used with Python program. The experimental results show almost perfect head pose detection performance when the head pose angle ranges from 0 to 180 degrees with 45 degrees steps.

The following section describes related research works followed by the proposed method. Then experiments are described followed by conclusion with some discussions and remarks.

## II. RELATED RESEARCH WORK

Computer input just by sight, human eyes only require head pose detection and head pose angle estimation.

Communication aid and computer input system with human eyes only is proposed [11]. Meanwhile, computer input by human eyes only and its applications are presented [12]. On the other hand, electric wheelchair control with gaze detection and

<sup>1</sup> <http://dx.doi.org/10.1109/TITS.2010.2044241>

<sup>2</sup> <http://dlib.net/>

eye blinking is proposed [13] together with electric wheelchair control with gaze detection and eye blinking [14].

Computer input with human eyes only using two Purkinje images which works in a real time basis without calibration is proposed [15]. Meanwhile, a prototype of electric wheelchair control by eye only for paralyzed use is created [16].

Robot arm utilized having meal support system based on computer input by human eyes only is also proposed and developed [17]. Also, a prototype of electric wheelchair controlled by eyes only for paralyzed users is created [18].

Autonomous control of eye based electric wheelchair with obstacle avoidance and shortest path finding based on Dijkstra algorithm, is attempted [19]. Meantime, eye-based human-computer interaction allowing phoning, reading e-book/e-comic/e-learning is created [20] together with eye based electric wheelchair control system-I(eye) can control EWC (Electric Wheelchair) [21].

Evaluation of users' impact for using the proposed eye based HCI: Human-Computer Interaction with moving and fixed keyboard by using EEG signals is conducted [22] together with electric wheelchair controlled by human eyes only with obstacle avoidance [23]. Also, evaluation of users' impact for using the proposed eye based HCI with moving and fixed keyboard by using EEG (Electroencephalography) signals is proposed with experimental validations [24].

Electric wheelchair controlled by human eyes only with obstacle avoidance is proposed and created [25] together with eye based HCI, a new keyboard for improving accuracy and minimizing fatigue effect [26].

Moving keyboard for eye based HCI is proposed [27]. Also, eye-based domestic robot allowing patient to be self-services and communications remotely is proposed and created [28].

Method for psychological status estimation by gaze location monitoring using eye based HCI is created and proposed [29]. Meanwhile, method for psychological status monitoring with line-of-sight vector changes (Human eyes movements) detected with wearing glass is proposed [30].

Wearable computing system with input output devices based on eye based HCI allowing location-based web services is proposed and realized [31]. Meanwhile, speed and vibration performance as well as obstacle avoidance performance of electric wheelchair controlled by human eyes only is evaluated [32] together with speed and vibration performance as well as obstacle avoidance performance of electric wheelchair controlled by human eyes only [33].

Service robot with communication aid together with routing controlled by human eyes is created [34]. On the other hand, information collection service system by human eyes for disabled persons is proposed [35]. Meanwhile, relations between psychological status and eye movements are investigated [36].

Method for 3D image representation with reducing the number of frames based on characteristics of human eyes is proposed [37]. Also, error analysis of line-of-sight estimation

using Purkinje images for Eye-Based Human Computer Interaction: EBHCI is proposed [38].

Mobile phone operations using human eyes only and its applications are created [39]. Meanwhile, method for thermal pain level prediction with eye motion using Support Vector Machine: SVM is proposed [40]. On the other hand, pedestrian safety with eye contact between autonomous car and pedestrian is proposed [41].

### III. 3D HEAD POSITION AND POSE MODEL

#### A. 3D Head Pose Model

In 3D head pose model is used to convert 2D face features into 3D head pose. The face features such as eyes, eyebrows, nose, and mouth are used. Head are modeled into 3 planar: XY, XZ, and YZ planar. Head pose is shown as rotation degree value ( $\theta$ ) of each planar. By using this model, 3D head pose is expected can be able to calculate only using 2D image. Head pose result is shown as  $\theta(x, y, z)$ .

Fig. 1 show head pose model on each planar. On Fig. 1(a), it shows head pose model on XY planar which look at face on front side. Central axis is assumed on bellow of mouth. When head is rotate, face features position will move follow their rotation against central axis. The new position of face features determines as  $R_i < \theta_i$ , where R is face feature radial and  $\theta$  is rotation angle. On Fig. 1(b), it shows head pose model on XY planar which look at lateral view.

Assume one of face feature has location  $P(x, y)$ , radial R, initial angle  $\theta_0$ , and central axis O (0, 0). If  $P(x, y)$  rotate against central axis, new point  $P_1(x_1, y_1)$  will be obtained. Both values can obtain rotation angle value. Rotation angle is calculated using equation (1).

$$\theta_{xy} = \tan^{-1} \left[ \frac{x'}{y'} \right] - \tan^{-1} \left[ \frac{x}{y} \right] \quad (1)$$

The same way, we also can calculate it for YZ and XZ planar using rotation radial of head. By assume rotation radial value, rotation angle for each planar will be known.

$$\theta_{yz} = \sin^{-1} \left[ \frac{y'}{R} \right] - \sin^{-1} \left[ \frac{y}{R} \right] \quad (2)$$

$$\theta_{xz} = \sin^{-1} \left[ \frac{x'}{R} \right] - \sin^{-1} \left[ \frac{x}{R} \right] \quad (3)$$

In a real condition, all information is shown on pixel coordinate. Therefore central axis will has  $O(x, y)$  coordinate and face feature will have  $P_i(x, y)$ . We can directly convert from pixel coordinate into rotation angle based on equation (2).

$$\theta_{xy} = \tan^{-1} \left[ \frac{O_x - P_x'}{O_y - P_y'} \right] - \tan^{-1} \left[ \frac{O_x - P_x}{O_y - P_y} \right] \quad (4)$$

$$\theta_{yz} = \sin^{-1} \left[ \frac{O_y - P_y'}{R} \right] - \sin^{-1} \left[ \frac{O_y - P_y}{R} \right] \quad (5)$$

$$\theta_{xz} = \sin^{-1} \left[ \frac{P_x' - O_x'}{R} \right] - \sin^{-1} \left[ \frac{P_x - O_x}{R} \right] \quad (6)$$

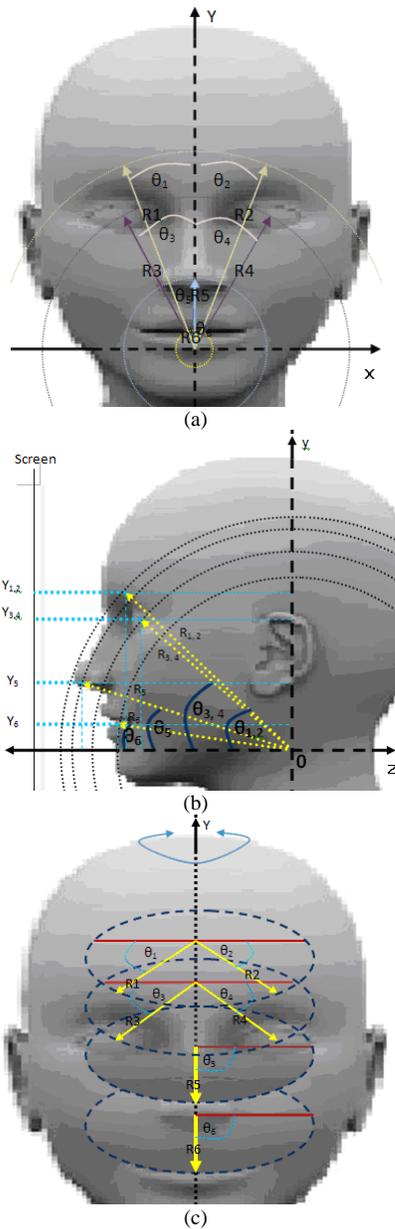


Fig. 1. 3D Head Model, (a) is Model on XY Planar, (b) is Model on YZ Planar, and (c) is Model on XZ Planar.

#### IV. PROPOSED METHOD

##### A. Head Pose Detection Overview

This section proposes a method for estimating the head posture of a pedestrian. Normally, on a road without a pedestrian crossing, when a pedestrian and a car driven by a person are close to each other, they exchange their intentions through nonverbal communication to ensure safe and secure traffic. However, in the case of an autonomous vehicle in which a person does not intervene in the operation, communication cannot be performed, and the following behavior cannot be assumed. As a result, it is easy to imagine that many people will be worried about autonomous vehicles running on the road and will feel uneasy when crossing the road.

OpenCV is an open-source computer vision and machine learning software library. It has C / C++, Python, Java, and MATLAB interfaces, and supports Windows, Linux, Android, and Mac OS. The library has over 2500 computer vision and machine learning algorithms. These are face detection and recognition, object identification, classification of human behavior in video, camera movement tracking, moving object tracking, 3D model extraction of objects, 3D point group generation from stereo cameras, image composition, scenes. It can be used to generate an entire high-resolution image, search for similar images from an image database, remove red eyes from images using a flash, and track eye movements. OpenCV is widely used by businesses, research groups and government agencies.

Dlib is an open-source software library written in C++. It is used in a wide range of fields such as robotics, embedded devices, mobile phones, and large-scale high-performance computing environments. In recent years, components for processing in a wide range of fields such as GUI (Graphical User Interface), machine learning, image processing, data mining, mathematical optimization, and Bayesian networks have been developed.

##### B. Representation of 3D Objects

A three-dimensional object with respect to a camera can be represented by the following two actions,

- 1) Translation: Moving the camera from one 3D position  $(X, Y, Z)$  to a new 3D position  $(X', Y', Z')$ . There are 3 degrees of freedom in movement, and it can move in the "X, Y, Z" directions.
- 2) Rotation: The camera can be rotated around the "X, Y, Z" axes. Rotation can be expressed by Euler angles (roll, pitch, yaw). In other words, it is possible to estimate the posture in three dimensions by finding three translations and rotations.

Fig. 2 shows an example of 3D representation of the face. Also, Fig. 3 shows the coordinate system conversion among the camera, the image (camera) and the world coordinate systems.

The coordinates of facial features shown in three dimensions are expressed in world coordinates. Three coordinate systems are used to estimate the attitude. If the attitude can be obtained, it will be possible to convert the 3D point in world coordinates to the 3D point in camera coordinates. The 3D points of the camera coordinates can be projected onto the image plane using camera-specific parameters such as focal length and lens distortion.



Fig. 2. 3D Mapping of the Face.

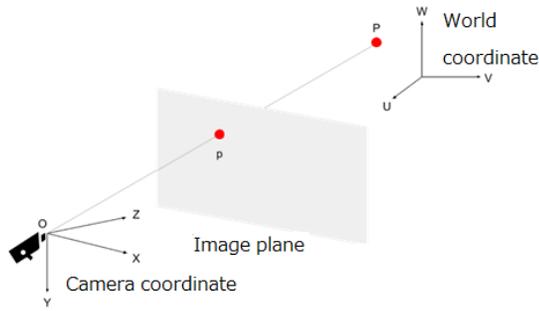


Fig. 3. Coordinate System Conversion.

Let the coordinate system fixed to the camera be  $(X, Y, Z)$ , the coordinate system fixed to the human head be  $(U, V, W)$ ,  $R$  be the rotation matrix, and  $t$  be the translation vector. The point  $P$  seen from the coordinate system fixed to the camera is expressed as follows.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} U \\ V \\ W \end{bmatrix} + t \quad (7)$$

Expressing this as an in-order transformation matrix, Eq. (8) is obtained.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [R|t] \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} \quad (8)$$

If the camera-specific parameters are known and the scale factor is  $s$ , then Eq. (9) is obtained.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = s \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (9)$$

The relationship between  $(x, y)$  and  $(X, Y, Z)$  can be expressed. Using this relationship, if  $R$  and  $t$  can be derived so that the error between the point  $p'$  that will be projected on the two-dimensional plane and the point  $p$  that is actually projected can be minimized, the attitude estimation will be performed.

### C. Head Pose Detection based on CNN

In the head posture estimation based on deep learning in image recognition, the angle of the face is divided into 45 degrees in the left-right direction as shown in Fig. 4, and the face faces at 90 degrees, 45 degrees, 0 degrees, -45 degrees, and -90 degrees. Also, Fig. 5 shows the definition of the head pose angle (the geometric relation between the car and the pedestrian).

Preliminary experiments have shown that it is difficult to discriminate even the finest angles of the face. For example, even if two images with only one degree difference in face angle are given, it is difficult to distinguish them because there is no difference in features between the two images. Moreover, in this study, it is only necessary to be able to grasp a rough angle, so we could not find the need to discriminate even a fine angle. Therefore, classification is performed in 5 classes of 90 degrees, 45 degrees, 0 degrees, -45 degrees, and -90 degrees.



Fig. 4. Examples of Head Pose with the different Pose Angles.



Fig. 5. Definition of the Head Pose Angle (the Geometric Relation between the Car and the Pedestrian).

The angle is based on the line of sight between the car and the pedestrian. The model used is a convolutional neural network (CNN). CNN is a forward propagation network that includes a convolution layer and a pooling layer. A learning method often used for image recognition and natural language processing.

### D. Face Detection with Dlib

A trained model distributed by Dlib is used to detect facial feature points. As shown in Fig. 6, 68 feature points could be detected, and 6 of them (nose tip, chin, left end of left eye, right end of right eye, left corner of mouth, right corner of mouth) were used for head posture estimation.

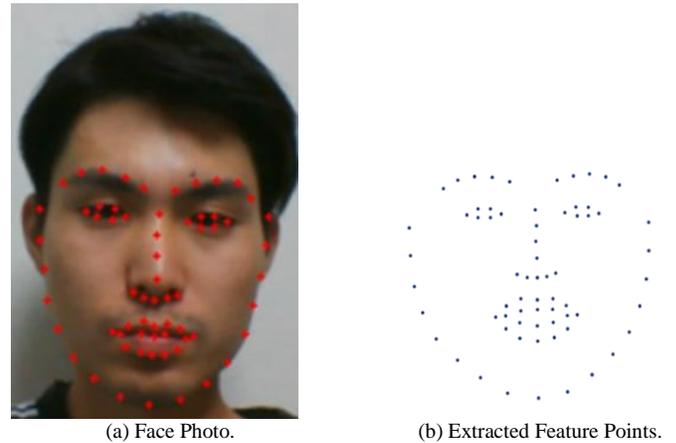


Fig. 6. Face Detection with Dlib.

## V. EXPERIMENTS

### A. Head Pose Detection

The camera used in the experiment was an HD (High Definition) webcam (manufactured by Sony Corporation) equipped with an "Exmor R for PC" CMOS (Complementary Metal Oxide Semiconductor) sensor, and the frame rate was 15 [fps]. The PC specifications for running this program are OS: Windows10 (64bit), CPU: Intel Core i5-5275U, and memory: 8.00GB.

The state of estimation is as shown in the figure below. Fig. 7 shows the coordinate system fixed to the head, which uses the rotation of the three axes to represent the angle. In addition, the estimation result expresses a three-dimensional view of the face with a green cube so that it is easy to understand visually, and as the estimation result, the angle of the face can be calculated accurately to a fine value. It was a good result.

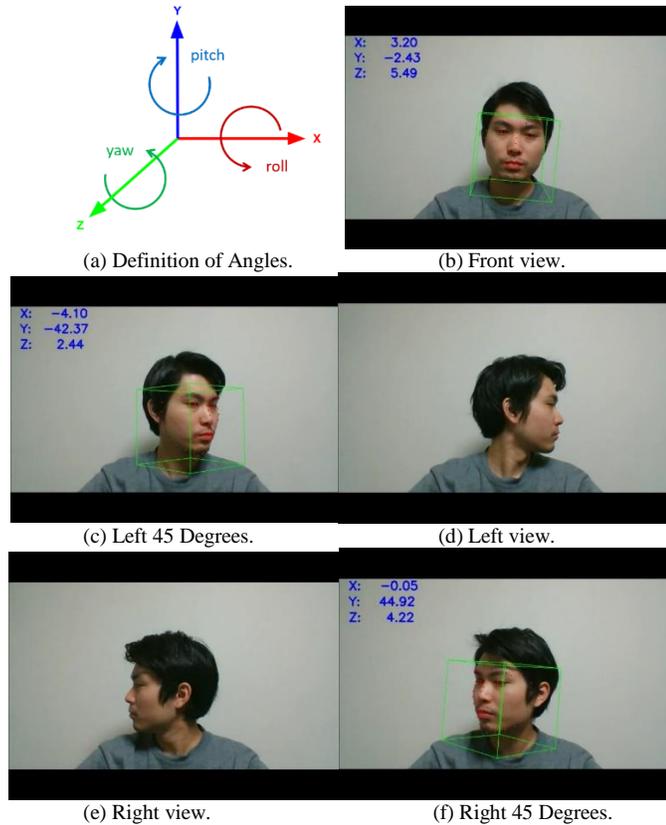


Fig. 7. Results from Head Pose Angle Detection.

### B. Head Pose Angle Detection

The verification experiment of head posture estimation and its result are described. The image used was Head Pose Image Database [23]. In the experiment, the data used was divided

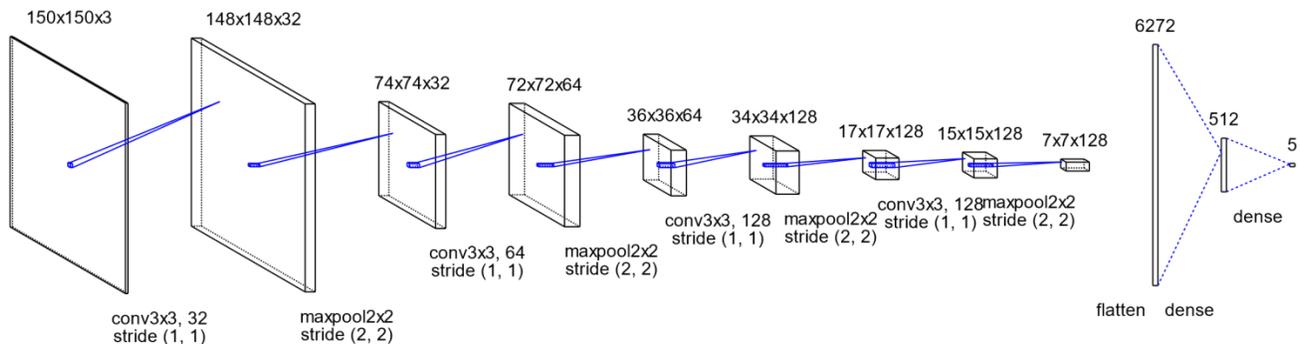


Fig. 9. CNN used.

into four patterns (RGB, RGB + thermography, grayscale image, RGB image assuming an image obtained by near infrared rays, and only the red channel element was extracted), and the accuracy was compared. Classify into five classes of 90 degrees, 45 degrees, 0 degrees, -45 degrees, and -90 degrees, respectively. The evaluation criteria for the classification results are set as shown in Fig. 8.

In the Fig. 8, marks are shown as follows:

⊙: Angle estimation was successful. The only thing left is how to respond (deceleration, warning).

○: Angle estimation failed, but within the permissible range. Correspondence is the same as ⊙ and there is no problem.

△: Angle estimation failed. There is a car approaching the pedestrian's field of view, but it may not be recognized, so safety first (not visible) is taken into consideration. Therefore, there is no problem in the end.

□: Angle estimation failed. However, as with △, safety first is taken as a result, so there is no problem in the end.

×: Angle estimation failed. Even though there are no approaching cars in the field of view of pedestrians, there is a possibility of taking dangerous measures.

Designated Angle	90°	⊙	○	△	□	□
	45°	○	⊙	△	□	□
	0°	△	△	⊙	□	□
	-45°	×	×	□	⊙	○
	-90°	×	×	□	○	⊙
	Estimated Angle	90°	45°	0°	-45°	-90°

Fig. 8. Evaluation Criteria.

In this study, we do not consider the actual response method that the automobile side will take after the estimation. The structure of the convolutional neural network was generated using an open-source Python implementation known as ConvNet Drawer<sup>3</sup>. Fig. 9 shows the CNN structure used.

C. Head Pose Detection Performance

There are three cameras, visible, NIR (Near Infrared) and thermal cameras for acquisition of face images. The purpose of this study is to detect head pose in all weather condition and in day and night-time condition. Therefore, the aforementioned three cameras are considered. As for the number of images of training data for learning process of CNN and the number of images for performance evaluation, Table I shows the numbers for each designated head pose angle for visible, NIR and thermal cameras while, Table II shows these numbers for visible and thermal cameras. In the later case, both of visible and thermal camera data are used together for training and performance evaluation.

The results using RGB images are shown in Fig. 10(a). The number of data used was 480, and the ratio of training data to verification data was 4:1. The facial images used for each class are data for 15 people. Most of the classified results were successful angle estimation for RGB images. Although there are parts where estimation fails in three places, we were able to create a model with high accuracy and no problems because it corresponds to the evaluation standard  $\bigcirc$ .

The number of data used was 480 RGB images and 160 thermography. The ratio of RGB images of each class to thermography is 3:1.

TABLE I. THE NUMBER OF IMAGES OF TRAINING DATA FOR LEARNING PROCESS OF CNN AND THE NUMBER OF IMAGES FOR PERFORMANCE EVALUATION FOR VISIBLE, NIR AND THERMAL CAMERAS

Head pose angle	No. of training data	No. of test data	Total
90°	72	18	90
45°	72	18	90
0°	90	30	120
-45°	72	18	90
-90°	72	18	90
Total	378	102	480

TABLE II. THE NUMBER OF IMAGES OF TRAINING DATA FOR LEARNING PROCESS OF CNN AND THE NUMBER OF IMAGES FOR PERFORMANCE EVALUATION FOR BOTH VISIBLE AND THERMAL CAMERAS

Head pose angle	No. of training data	No. of test data	Total
90°	96	24	120
45°	96	24	120
0°	120	30	160
-45°	96	24	120
-90°	96	24	120
Total	512	128	640

In the classification results, erroneous estimation was found in the parts corresponding to  $\triangle$  and  $\square$ . Although there is no erroneous estimation at the point x, it can be said that the accuracy is lower than the result of only RGB images, which are all within the permissible range. In addition, although it was within the permissible range, there were many false estimates at 90 degrees and 45 degrees as shown in Fig. 10(b).

The data used is a grayscale version of all 480 images used in the experiment using only RGB images.

The results were predicted to some extent, but they were the worst compared to the experimental results of the other three patterns as shown in Fig. 10(c).

This is thought to be due to the fact that grayscale images have less information than RGB images.

As the amount of information is reduced, it becomes difficult to capture the features. The results of using only the image obtained by extracting only the red channel elements from the RGB image assuming the image obtained by near infrared rays are described.

The data used is only the red channel elements extracted from all 480 images used in the experiment using only RGB images. Originally, the image actually taken by the near-infrared camera should be used, but since there is no equipment, the simulation was performed in this way. By extracting only, the elements of the red channel from the RGB image, information on short wavelengths can be dropped, so we decided to regard the image used this time as an image obtained by shooting with a near-infrared camera.

Most of the classification results were successful angle estimation. Although there was an erroneous estimation outside the permissible range at only one location, the accuracy was second only to the result using only RGB images as shown in Fig. 10(d).

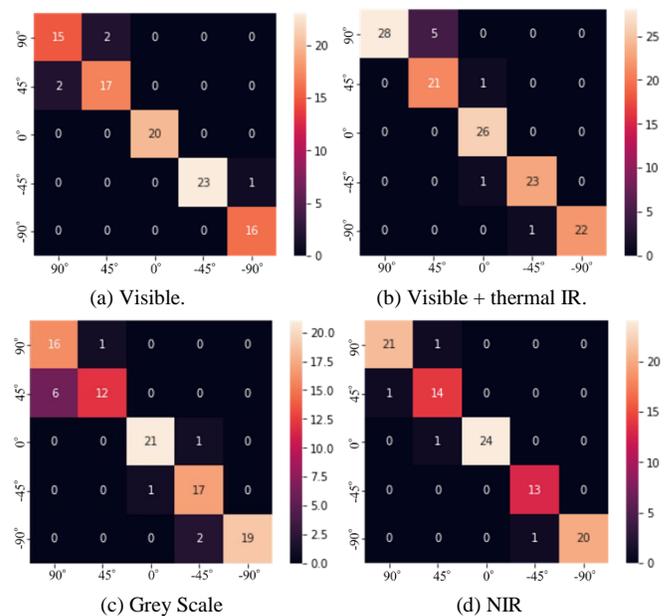


Fig. 10. Results from the Head Pose Detection Performance.

<sup>3</sup> <https://github.com/yu4u/convnet-drawer>,

In the head posture estimation using the feature points, even if the machine used has low specifications, the program can be processed with almost no delay, and the angle can be calculated with high accuracy. However, it did not work when the feature points were turned in a hidden direction, resulting in a disappointing result. If the angle can be calculated using the feature points on only one side of the face, it may be possible to handle it even when facing sideways.

In head posture estimation using deep learning, we conducted experiments with four patterns and classified angles. The order of accuracy was 1) RGB image, 2) Near infrared ray (red channel), 3) RGB image and thermography, and 4) Gray scale. From this result, it is considered better to use a normal camera for estimation in bright hours such as morning and noon. At night or in poor visibility, it is better to irradiate the front of the vehicle with near-infrared rays to make estimation, although the accuracy will be slightly lower. In addition, since automobiles usually illuminate headlights at night, we think that it may be possible to clearly capture pedestrians in combination with near infrared rays.

## VI. CONCLUSION

In this study, we proposed a method for estimating the head posture of a pedestrian and conducted a verification experiment in order for the autonomous vehicle to make contact with the pedestrian. Although the facial feature points could be detected in detail by the method using facial feature points, the necessary feature points could not be extracted when facing sideways, so angle classification was performed using a convolutional neural network.

The images used were RGB images, RGB images + thermography, grayscale images, and RGB images assuming images obtained by near infrared rays, with only the red channel elements extracted.

As a result, the RGB image model was the most accurate, but considering the criteria set, the RGB image model was used for morning and daytime detection, and the near-infrared image was used for nighttime and rainy weather scenes. It turned out that it is better to use the model obtained by the training in.

## VII. FUTURE RESEARCH WORKS

As a future task, we would like to create a mechanism to feed back the cognitive status to pedestrians. In addition, we would like to verify whether this system can behave in the same way as a human-driven vehicle and a pedestrian when this system is installed in an actual autonomous driving vehicle.

## ACKNOWLEDGMENT

The author would like to thank Professor Dr. Osamu Fukuda for their valuable discussions.

## REFERENCES

- [1] A. Al-Rahayfeh and M. Faezipour, "Eye Tracking and Head Movement Detection: A State-of-Art Survey," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 1, pp. 11- 22, 2013. <http://dx.doi.org/10.1109/JTEHM.2013.2289879>, 2013.
- [2] K. Liu, Y. P. Luo, G. Tei, and S. Y. Yang, "Attention recognition of drivers based on E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness" *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 2, pp. 300-311, Jun. 2010.
- [3] D. J. Kupetz, S. A. Wentzell, and B. F. BuSha, "Head motion controlled power wheelchair" in *Proc. IEEE 36th Annu. Northeast Bioeng. Conf.*, Mar. 2010, pp. 1-2, 2010.
- [4] L. M. King, H. T. Nguyen, and P. B. Taylor, "Hands-free headmovement gesture recognition using artificial neural networks and the magnified gradient function," in *Proc. 27th Annu. Conf. Eng. Med. Biol.*, 2005, pp. 2063-2066. <http://dx.doi.org/10.1109/ieems.2005.1616864>, 2005.
- [5] S. T. Nguyen, H. T. Nguyen, P. B. Taylor, and J. Middleton, "Improved head direction command classification using an optimised Bayesian neural network" in *Proc. 28th Annu. Int. Conf. EMBS*, 2006, pp. 5679-5682, 2006.
- [6] S. Manogna, S. Vaishnavi, and B. Geethanjali, "Head movement based assist system for physically challenged" in *Proc. 4<sup>th</sup> ICBBE*, 2010, pp. 1-4, 2010.
- [7] S. Kim, M. Park, S. Anumas, and J. Yoo, "Head mouse system based on gyro- and opto-sensors" in *Proc. 3rd Int. Conf. BMEI*, vol. 4, 2010, pp. 1503-1506, 2010.
- [8] K. Satoh, S. Uchiyama, and H. Yamamoto, "A head tracking method using bird's-eye view camera and gyroscope" in *Proc. 3<sup>rd</sup> IEEE/ACM ISMAR*, Nov. 2004, pp. 202-211, 2004.
- [9] Kohei Arai, H. Tolle, A. Serita, "Mobile Devices Based 3D Image Display Depending on User's Actions and Movements. *International Journal of Advanced Research in Artificial Intelligence (IJARAI)*, 2013, vol 2, no. 6. pp.71-78, 2013.
- [10] H. Tolle, A. Pinandito, EM. Adams J., K. Arai, "Virtual reality game controlled with user's head and body movement detection using smartphone sensors". *ARPN Journal of Engineering and Applied Sciences*. Nov. 2015, vol 10, no 20, pp 9776-9782, 2015.
- [11] Kohei Arai and Kenro Yajima, *Communication Aid and Computer Input System with Human Eyes Only*, Electronics and Communications in Japan, Volume 93, Number 12, 2010, pages 1-9, John Wiley and Sons, Inc., 2010.
- [12] Kohei Arai, *Computer Input by Human Eyes Only and Its Applications, Intelligent Systems in Science and Information*, 2014, Studies in Computer Intelligence, 591, 1-22, Springer Publishing Co. Ltd., 2015.
- [13] Djoko Purwanto, Ronny Mardiyanto, Kohei Arai, Electric wheel chair control with gaze detection and eye blinking, *Proceedings of the International Symposium on Artificial Life and Robotics*, GS9-4, 2009.
- [14] Djoko Purwanto, Ronny Mardiyanto and Kohei Arai, Electric wheel chair control with gaze detection and eye blinking, *Artificial Life and Robotics*, AROB Journal, 14, 694,397-400, 2009.
- [15] Kohei Arai and Makoto Yamaura, Computer input with human eyes only using two Purkinje images which works in a real time basis without calibration, *International Journal of Human Computer Interaction*, 1,3, 71-82,2010.
- [16] Kohei Arai, Ronny Mardiyanto, A prototype of electric wheelchair control by eye only for paralyzed user, *Journal of Robotics and Mechatronics*, 23, 1, 66-75, 2010.
- [17] Kohei Arai, Kenro Yajima, Robot arm utilized having meal support system based on computer input by human eyes only, *International Journal of Human-Computer Interaction*, 2, 1, 120-128, 2011.
- [18] Kohei Arai and Ronny Mardiyanto, A prototype of electric wheel chair controlled by eyes only for paralyzed users, *Journal of Robotics and Mechatronics*, 23, 1, 66-75, 2011.
- [19] Kohei Arai, Ronny Mardiyanto, Autonomous control of eye based electric wheel chair with obstacle avoidance and shortest path finding based on Dijkstra algorithm, *International Journal of Advanced Computer Science and Applications*, 2, 12, 19-25, 2011.
- [20] Kohei Arai, Ronny Mardiyanto, Eye-based human-computer interaction allowing phoning, reading e-book/e-comic/e-learning, Internet browsing and TV information extraction, *International Journal of Advanced Computer Science and Applications*, 2, 12, 26-32, 2011.
- [21] Kohei Arai, Ronny Mardiyanto, Eye based electric wheel chair control system-I(eye) can control EWC-, *International Journal of Advanced Computer Science and Applications*, 2, 12, 98-105, 2011.

- [22] Kohei Arai, Ronny Mardiyanto, Evaluation of users' impact for using the proposed eye based HCI with moving and fixed keyboard by using eeg signals, *International Journal of Research and Reviews on Computer Science*, 2, 6, 1228-1234, 2011.
- [23] Kohei Arai, Ronny Mardiyanto, Electric wheel chair controlled by human eyes only with obstacle avoidance, *International Journal of Research and Reviews on Computer Science*, 2, 6, 1235-1242, 2011.
- [24] Kohei Arai, R.Mardiyanto, Evaluation of users' impact for using the proposed eye based HCI with moving and fixed keyboard by using eeg signals, *International Journal of Research and review on Computer Science*, 2, 6, 1228-1234, 2012.
- [25] Kohei Arai, R.Mardiyanto, Electric wheel chair controlled by human eyes only with obstacle avoidance, *International Journal of Research and review on Computer Science*, 2, 6, 1235-1242, 2012.
- [26] R.Mardiyanto, Kohei Arai, Eye-based Human Computer Interaction (HCI) A new keyboard for improving accuracy and minimizing fatigue effect, *Scientific Journal Kursor*, (ISSN 0216-0544), 6, 3, 1-4, 2012.
- [27] Kohei Arai, R.Mardiyanto, Moving keyboard for eye-based Human Computer Interaction: HCI, *Journal of Image and Electronics Society of Japan*, 41, 4, 398-405, 2012.
- [28] Kohei Arai, Ronny Mardiyanto, Eye-based domestic robot allowing patient to be self-services and communications remotely, *International Journal of Advanced Research in Artificial Intelligence*, 2, 2, 29-33, 2013.
- [29] Kohei Arai, Ronny Mardiyanto, Method for psychological status estimation by gaze location monitoring using eye-based Human-Computer Interaction, *International Journal of Advanced Computer Science and Applications*, 4, 3, 199-206, 2013.
- [30] Kohei Arai, Kiyoshi Hasegawa, Method for psychological status monitoring with line of sight vector changes (Human eyes movements) detected with wearing glass, *International Journal of Advanced Research in Artificial Intelligence*, 2, 6, 65-70, 2013.
- [31] Kohei Arai, Wearable computing system with input output devices based on eye-based Human Computer Interaction: HCI allowing location based web services, *International Journal of Advanced Research in Artificial Intelligence*, 2, 8, 34-39, 2013.
- [32] Kohei Arai Ronny Mardiyanto, Speed and vibration performance as well as obstacle avoidance performance of electric wheel chair controlled by human eyes only, *International Journal of Advanced Research in Artificial Intelligence*, 3, 1, 8-15, 2014.
- [33] Kohei Arai Ronny Mardiyanto, Speed and vibration performance as well as obstacle avoidance performance of electric wheel chair controlled by human eyes only, *International Journal of Advanced Research in Artificial Intelligence*, 3, 1, 8-15, 2014.
- [34] Kohei Arai, Service robot with communication aid together with routing controlled by human eyes, *Journal of Image Laboratory*, 25, 6, 24-29, 2014.
- [35] Kohei Arai, Information collection service system by human eyes for disable persons, *Journal of Image Laboratory*, 25, 11, 1-7, 2014.
- [36] Kohei Arai, Relations between psychological status and eye movements, *International Journal of Advanced Research on Artificial Intelligence*, 4, 6, 16-22, 2015.
- [37] Kohei Arai, Method for 3D Image Representation with Reducing the Number of Frames Based on Characteristics of Human Eyes, *International Journal of Advanced Research on Artificial Intelligence*, 5, 8, 7-12, 2016.
- [38] Kohei Arai, Error Analysis of Line of Sight Estimation Using Purkinje Images for Eye-Based Human Computer Interaction: EBHCI, *International Journal of Advanced Research on Artificial Intelligence*, 5, 10, 14-23, 2016.
- [39] Kohei Arai, Mobile Phone Operations using Human Eyes Only and Its Applications, *International Journal of Advanced Computer Science and Applications IJACSA*, 9, 3, 2018.
- [40] Kohei Arai, Method for Thermal Pain Level Prediction with Eye Motion using SVM, *International Journal of Advanced Computer Science and Applications IJACSA*, 9, 4, 170-175, 2018.
- [41] Kohei Arai, Akihiro Yamashita, Hiroshi Okumura, Pedestrian safety with eye contact between autonomous car and pedestrian, *International Journal of Advanced Computer Science and Applications IJACSA*, 10, 5, 161-165, 2019.

#### AUTHOR'S PROFILE

**Kohei Arai**, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is a Science Council of Japan Special Member since 2012. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Science Commission "A" of ICSU/COSPAR since 2008 then he is now award committee member of ICSU/COSPAR. He wrote 55 books and published 620 journal papers as well as 450 conference papers. He received 66 of awards including ICSU/COSPAR Vikram Sarabhai Medal in 2016, and Science award of Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. <http://teagis.ip.is.saga-u.ac.jp/index.html>.