Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

A Deep Neural Network Based Human Following Robot with Fuzzy Control

Yin Yin Aye, Kyaw Thiha, and Mi Mi Myint Pyu
Department of Mechatronic Engineering
Mandalay Technological University
Mandalay, Myanmar
{yinyinaye, kyawthiha, mimimyintpyu}@mtu.edu.mm

Keigo Watanabe
Department of Intelligent Mechanical Systems
Okayama University
3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530, Japan
watanabe@sys.okayama-u.ac.jp

***Abstract*— This research aims to build a human following robot system based on a deep neural network algorithm in which a fuzzy controller controls the robot velocity and keeps the target person in the centre position of robot's view. Firstly, the system utilizes the deep neural network algorithm to detect a target person in the video sequence captured from a real sense D435 depth camera mounted on the mobile robot. Then, the system calculates the centre position of the target human and acquires the depth value of target human. Finally, these data are used as the inputs of a fuzzy controller to control the velocity and steering of the robot during tracking. Especially, the velocity of the robot which is normally limited as a constant in most existing human following robot systems is controlled by a fuzzy controller in this paper. The proposed system is verified through the experiments for a four-wheel steered mobile robot.**

***Index Terms*— Human tracking, Deep neural network, Fuzzy control, A car-like mobile robot**

I. Introduction

In current decade, service robots which can operate in an unstructured environment have attracted a great deal of attention from research organization. In applications such as surveillance, intelligent transportation, human robot interaction and so on, it needs to build the service robots. An important requirement for a service robot is that the robot is able to follow a target human in homes or work places. Although a number of different approaches have been developed under this area, to find a robust method for human tracking robots in real time is still a rough task.

Most of the recently developed approaches on human-tracking robot are limited to an operating environment and require humans to wear sensors or color clothes. A human following system used a 2D camera to detect and track the human who carried the LED lights, was proposed as early research. To detect the target person, several approaches such as distribution based on skin color, motion-based, and machine learning have been developed in [1], [2], and [3], respectively. On the other hand, laser based human tracking and both a camera and a laser fusion tacking approaches have been described in [4] and [5]. When the human tracking algorithm is implemented in the outdoor environment, the data from sensors may have more noise and also the target human may be lost temporarily. Therefore, a human tracking system will be an enormously difficult task in such unstructured environments. In [6], a Particle Filter is introduced to detect human. However, the single vision-sensor methods only detect the orientation of human and do not provide depth information like a stereo-visual sensor.

Compared to the conventional machine learning techniques, the structure of deep neural networks (DNN) involves multiple hidden layers to enable the extraction of features that are deeply embedded in the data, forming abstract concepts in a hierarchy manner. DNN consists of hierarchically arranged trainable layers that automatically detect and learn patterns underlying the images. In OpenCV, it includes a deep learning module with support for a number of deep learning frameworks such as Caffe, Tensor Flow and so on. OpenCV DNN API can detect people using a pre-trained Caffe model (MobileNet +SSD). The applications of 2D computer vision have been limited by the lack of an important depth dimension. The 3D depth sensing in real time has long been a major challenge for the robotics industry these days. 3D vision enables machines to accurately understand shapes, sizes, distances and to maneuver in the real 3D world. In this research, OpenCV DNN API is utilized to detect a target person in the video sequence captured from a real sense 3D camera mounted on the mobile robot. Our research aims to build a novel human tracking system based on a DNN algorithm.

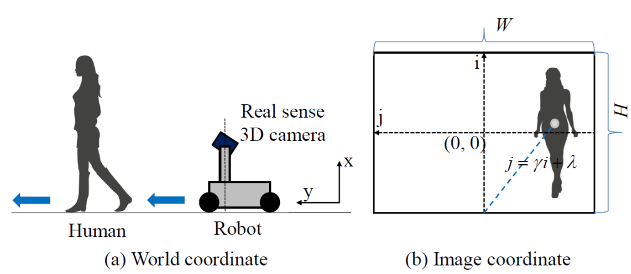

Fig. 1. Concept of a human following robot

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

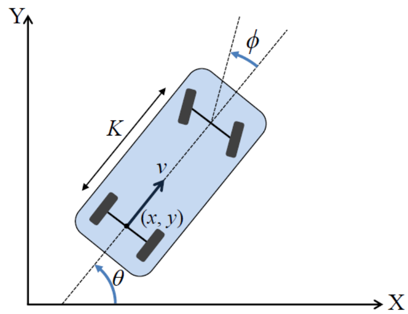

Fig. 2. Model of a mobile robot

In this work, the image-based control method (IBCM) is selected as a suitable one for tracking a target human by a robot. The 3D model of the working environment is not needed to utilize in the IBCM. Moreover, using this method, the robot is able to follow the human by controlling only image information which is acquired from a camera image without using the robot position. IBCM acts like human that it does not refer to its own position. It has advantages on reductions the computational time and the camera calibration process [7]. Aye *et al*. [8, 9] have already studied a forward parking problem without cutting the wheel by such an image based control approach, where the effectiveness was shown in several simulations and experiments with an actual setup. In this research, adopting a fuzzy controller by [10] and [11], allows us to design it for tracking the target human in an appropriate center position of the robot's view, as well as controlling the time-varying robot velocities during tracking. The experiments confirm that the designed fuzzy controller is able to track the human with time-varying velocity.

This paper describes the problem setting in Sect. 2 and the generation of a target line for a human following system in Sect. 3. In Sect. 4 the fuzzy controller is designed for a nonholonomic mobile robot with time varying velocity. Some experimental results are given to demonstrate the effectiveness of the proposed method in Sect. 5. Consideration and conclusion are summarized in Sects. 5 and 6, respectively.

II. Problem Setting

The problem addressed in this research is to build a human following robot using a fuzzy logic controller. Figure 1 shows the schematic concept of the present method. As shown in the figure, x-y and i-j coordinates are utilized to define the world coordinate and image coordinate for the camera where the image size, W (width) $\times$ H (height) is measured in pixel. A four-wheel mobile robot shown in Fig. 2 is used as a controlled object in this research. The kinematic model is derived under the assumption of rolling without slippage, on horizontal ground and it is represented by:

$$\begin{aligned} \dot{x} &= v\cos\theta \\ \dot{y} &= v\sin\theta \\ \dot{\theta} &= \frac{\tan\phi}{K} \end{aligned} \tag{1}$$

The position and orientation of the vehicle are described as (x, y, θ) and the wheelbase is denoted by K. To track the target human, the motion control of this robot is achieved by dealing with the steering angle denoted as ϕ, and the forward velocity denoted as v.

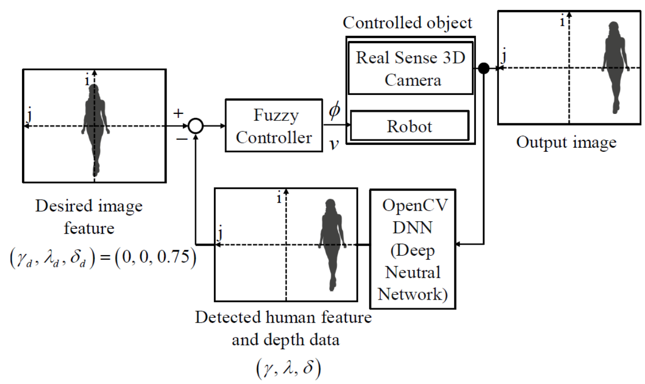

Fig. 3. Block diagram of the proposed human tracking system

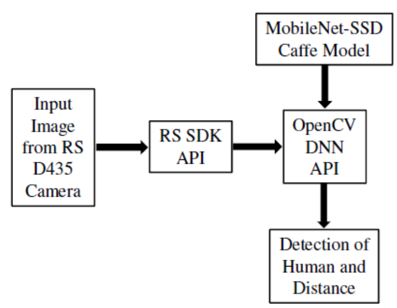

Fig. 4. Schematic of a human detection system using a DNN

The block diagram of the proposed human tracking control system is shown in Fig. 3. It mainly consists of two parts. In the first part, a deep neural network algorithm is utilized to detect a target person in the video sequence captured from the real sense 3D camera mounted on the mobile robot. The position and orientation of target human is calculated in order to generate the time-varying target line for human tracking system. Then, the slope and intercept of the target line, and the depth value obtained from the real sense depth camera are sent to the fuzzy logic controller as the control inputs. Note that, the minimum human detection distance is about 0.75 [m] away from the 3D camera mounted on the mobile robot. In the second part, a fuzzy controller is designed to follow the target person by the robot with time-varying velocity. The control inputs, steering angle and forward velocity, of our human tracking system are determined according to a model-free fuzzy tracking control.

III. Generation of a Target Line for a Human Following System

A. Human Detection Using a Deep Neural Network

In this research, a DNN algorithm is utilized to detect a target person in the video sequence captured from a real sense depth camera mounted on the mobile robot. The power of DNN is exploited for the problem of human detection in which the localization of target human is precisely tried. Figure 4 shows the schematic diagram of a human detection system using a DNN algorithm. As shown in Fig. 4, the color and depth frames captured from the 3D camera are converted to OpenCV Matrices.

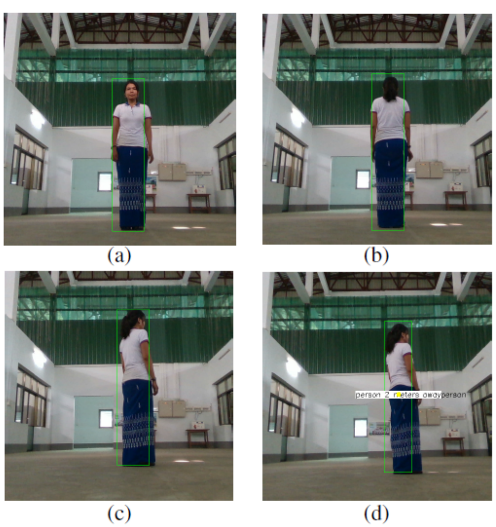

Fig. 5. Detected human (a) front view, (b) back view, (c) side view, and (d) with depth value and center position

Then, DNN API working with pre-trained deep learning models is applied to detect a human and calculate a distance between the detected human and the robot. The front, back, and side views of human detection results are shown in Fig. 5 (a), (b), and (c). The height of detected person is determined by using a combination of RGB data that are images from the camera and depth information. And also, the centre position of the target human is calculated based on the bounding box of target human and it is shown as yellow point in Fig. 5 (d). The distance between the detected human and the robot is mentioned in a small white rectangle of 2D image as well as in a 3D image as shown in Fig. 5 (d) and Fig. 6.

B. Calculation of Control Inputs for a Fuzzy Controller

To calculate the control inputs and generate the target line for a human tracking system, as first, the world coordinates (x, y) axis is changed to the image coordinates (j, i) axis by using the following equation:

$$\begin{aligned} j &= \frac{W}{2} - \mathrm{x} \\ i &= \frac{H}{2} - \mathrm{y} \end{aligned} \tag{2}$$

Figure 7 shows the changing of the world coordinate to image coordinate. As shown in Fig. 7, the center position of the detected person and the camera center position are denoted as (i_h, j_h) and (i_c, j_c). The time-varying target line for human tracking system is generated by using these data. The red line is the generated target line of the human tracking system. The slope, γ and intercept, λ are calculated using the following equation:

$$\begin{aligned} \gamma &= \frac{-j_h}{i_c - i_h} \\ \lambda &= \frac{i_c j_h - i_h}{i_c - i_h} \end{aligned} \tag{3}$$

The slope and intercept of the target line and the depth value of detected human are given to the fuzzy controller as the control inputs.

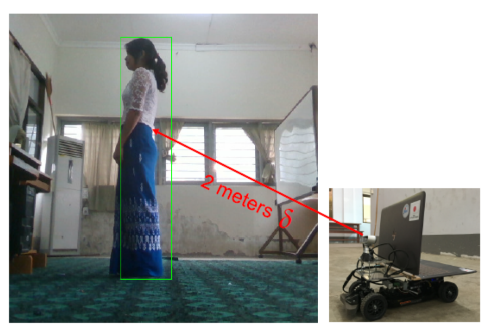

Fig. 6. Distance measurement between the robot and human

IV. Design of a Fuzzy Controller

The fuzzy controller designed to track the target human by the robot with time-varying velocity is described in this section. To construct the fuzzy controller, three consecutive steps are required such as fuzzification of state variables, calculation of grade of each rule and defuzzification of input values.

Fuzzification is the process of converting the real numbers to the linguistic representations with grades. The widths of the membership functions are determined by performing some preliminary experiments depending on the properties of the 3D camera and the robot. It is decided to utilize the triangular membership functions for the input variables γ, λ and δ and the singleton membership functions for output variables v, ϕ in this research. The membership functions of each state variable are given in Fig. 8, where fuzzy term sets are denoted by NE (Negative), ZE (Zero), PE (Positive), PES (Positive Small), and PEB (Positive Big). The function, represents the membership functions to calculate the grades of states and the range of each membership function to be within [0, 1].

The fuzzy reasoning rules for the human tracking problem are summarized in Table I and Table II. Fuzzy rules for controlling the velocity and steering of the robot are constructed depending on the use of fuzzified state variables. The min-max centroid method is utilized to calculate the grade of each rule in this study. Finally, the defuzzification process is performed on the fuzzy outputs which are obtained as fuzzy sets in order to determine the control inputs of the robot as real numbers. The defuzzification strategy is the weighted average method, which is given by:

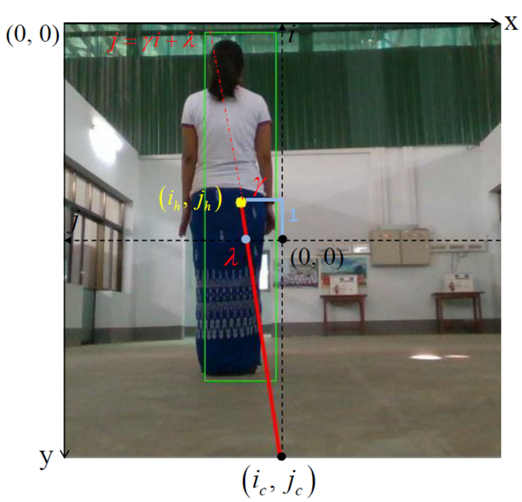

Fig. 7. Target line generated in an image

TABLE I
Fuzzy rules for ϕ

No.	γ	λ	ϕ
1	NE	NE	NE
2		ZE	NE
3		PE	ZE
4	ZE	NE	NE
5		ZE	ZE
6		PE	PE
7	PE	NE	ZE
8		ZE	PE
9		PE	PE

$$
\begin{aligned}
v &= \frac{0 \cdot h_{s,\mathrm{ZE}} + 0.5 \cdot h_{s,\mathrm{PES}} + 1.0 \cdot h_{s,\mathrm{PEB}}}{h_{s,\mathrm{ZE}} + h_{s,\mathrm{PES}} + h_{s,\mathrm{PEB}}} \\
\phi &= \frac{-0.349 \cdot h_{\phi,\mathrm{NE}} + 0 \cdot h_{\phi,\mathrm{ZE}} + 0.349 \cdot h_{\phi,\mathrm{PE}}}{h_{\phi,\mathrm{NE}} + h_{\phi,\mathrm{ZE}} + h_{\phi,\mathrm{PE}}}
\end{aligned}
\tag{4}
$$

Here, $h_{v,\mathrm{PES}}$ denotes the grade of conformity with the consequent "$v = \mathrm{PES}$." In our previous work [12], the fuzzy controller was designed to control only the steering angle of the robot and the distance between the robot and the target person was maintained about 0.75 m. In this paper, our controller is developed to control not only the steering angle but also the velocity of the robot which is usually assumed to be constant, as time varying one. The rules for the velocity control are constructed based on the distance between the target human and the robot, and the steering angle of the robot.

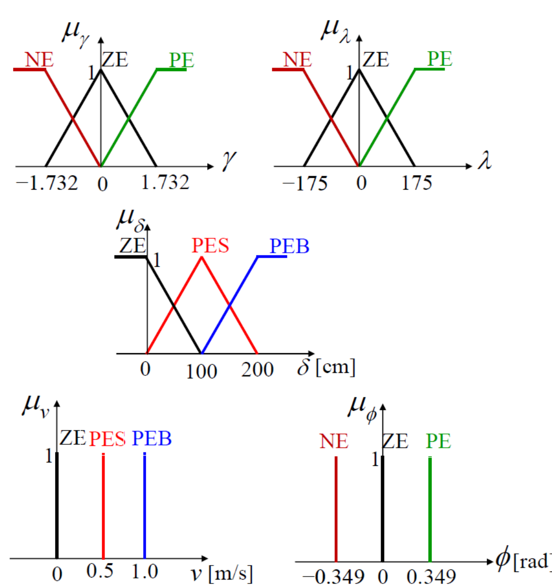

Fig. 8. Membership functions for human tracking

TABLE II

FUZZY RULES FOR v

No.	δ	ϕ	v
1	ZE		ZE
2	PES		PES
3	PEB	NE	PES
4		ZE	PEB
5		PE	PES

V. EXPERIMENTAL RESULTS

A system for automatically detecting the human based on a DNN and following human using a fuzzy controller is tested in this experiment. The robot utilized in this experiment was a four-wheel steered mobile robot. As shown in Fig. 9, the modified remote control car is used as the controlled object which is controlled by a microcontroller. The real sense D435 sensor is used to detect the human and acquire the distance in the unstructured environment. A mini laptop is utilized to extract the image parameters $(\gamma, \lambda, \delta)$ based on DNN and obtain the control inputs (v, ϕ) from the fuzzy controller. These control inputs are sent to the microcontroller, it controls the DC motor and the servomotor of the robot. DC motor is utilized to control the velocity and the servo motor is utilized to control the steering angle of the robot.

In this experiment, the detected human is tracked by the robot with time-varying velocity. The target person motion is to move forward, turn right and left, and finally to stop as shown in Fig. 10. The acquired distance and generated target line, red line, of our human tracking system are shown in Fig. 11. The control inputs, the velocity and steering angle values, are tested when the target human is moving forward, turning right and left and then stop. The recorded control inputs for this test are shown in Figs. 12 and 13, and the value of parameters on the captured images for the human tracking system is given in Fig. 14.

According to the recorded steering values as shown in Fig. 12, the robot can track the target human well. When the human moved forward, there was a constant steering value, whereas there was a negative steering value when the human turned to right. When the robot with maximum constant velocity tracks the target human, moving on curvilinear, the robot will be failed to track the human. In this research, the velocity control seems to exert its powerful effect on curvilinear human tracking. And the results proved that our approach was able to track the human by the robot with time-varying velocity.

VI. CONSIDERATIONS

The experimental results show the effectiveness of the proposed control system. Adding the velocity control in this human tracking system improved the turning performing of the robot. Especially, the reduction of velocity can protect the robot from failure tracking the human during the curvature turning. The robot followed the target human well, even if the velocity of human, moving on a straightforward, accelerated. On the other hand, the robot failed to track the human when the human body turned in a corner suddenly. It is attributed to the fact that the OpenCV DNN API utilized in this research can detect the human at least 0.75 [m] away from the 3D camera mounted on the mobile robot.

VII. CONCLUSIONS

In this paper, a novel human following robot system based on the DNN algorithm has been described. The objective was that a robot equipped a real sense D435 depth camera was to track a target human using a fuzzy controller. In this research, the fuzzy controller which operates a robot velocity and steering angle was proposed to use a time-varying robot velocity that is usually assumed to be constant. And 3D camera provided a low cost solution and better accuracy compared to the other sensor. The DNN based human detection method gave the better detection results.

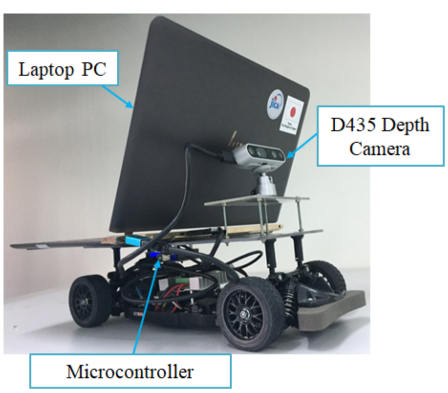

Fig. 9. Real robot for the experiment

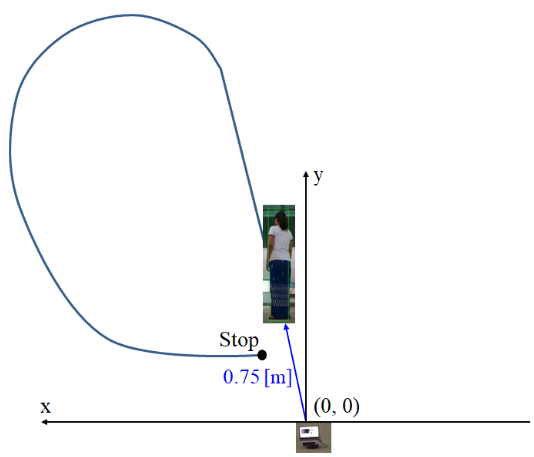

Fig. 10. Experimental environment

ACKNOWLEDGMENT

This research received funding from the Japan International Cooperation Agency (JICA) EEHT Project, which is sponsored by the Japanese government. I would like to thank JICA EEHE Project for financial support.

REFERENCES

[1] P. Vadakkepat, P. Lim, L. C. De Silva, L. Jing, and L. L. Ling, "Multimodal approach to human-face detection and tracking," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 3, pp. 1385–1393, March 2008.

[2] A. Utsumi and N. Tetsutani, "Human tracking using multiple camera based head appearance modeling," in *Proc. of IEEE International Conference on Automation Face and Gesture Recognition*, May 2004, pp. 657–662.

[3] P. Viola and M. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[4] R. Merwe, A. Doucet, N. Freitas, et al., "The unscented particle filter," in Technical Report CUED/F-INFENG/TR 380, Cambridge University Engineering Department, August 2000.

[5] Hough and V. C. Paul, "Method and means for recognizing complex patterns," US Patent No. 3,069,654, December 1962.

[6] M. Isard and A. Blake, "Visual tracking by stochastic propagation of conditional density," in *Proc. 4th European Conference Computer Vision*, Cambridge, UK, April 1996.

[7] B. Espiau, "Effect of camera calibration errors on visual servoing," *The 3rd international symposium on experimental robotics* , Japan, 1995, pp. 187–193.

[8] Y. Y. Aye, J. Ukida, K. Watanabe, S. Maeyama, and I. Nagai, "Design of an image-based fuzzy controller for autonomous parking of four-wheeled mobile robots," *International Journal of Applied Electromagnetics and Mechanics*, vol. 52, no. 34, pp. 859–865, 2016.

[9] Y. Y. Aye, K. Watanabe, S. Maeyama, and I. Nagai, "An automatic parking system using an optimized image-based fuzzy controller by genetic algorithms," *Artificial Life and Robotics*, vol. 22, no. 1, pp. 139–144, 2017.

[10] L. A. Zadeh, "Fuzzy algorithms," *Information and Control*, vol. 12, pp. 94–102, 1968.

[11] E. H. Mamdani, "Advances in the linguistic synthesis of of fuzzy controller," *International Journal of Man-Machine Studies*, vol. 8, no. 6, pp. 669–678, 1976.

[12] Y. Y. Aye, W. P. Ei, K. Watanabe, and S. Yadanar, "Design of a fuzzy controller for a human following robot system based on a deep neutral network," *The 2th International Conference on Science, Technology and Innovation, Mandalay by IEEE*, Myanmar, 2019, pp. 147–151.

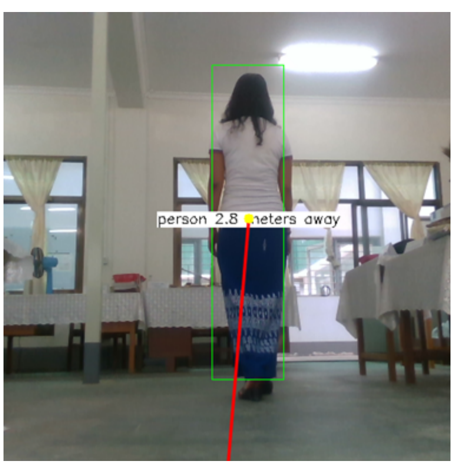

Fig. 11. Generated target line

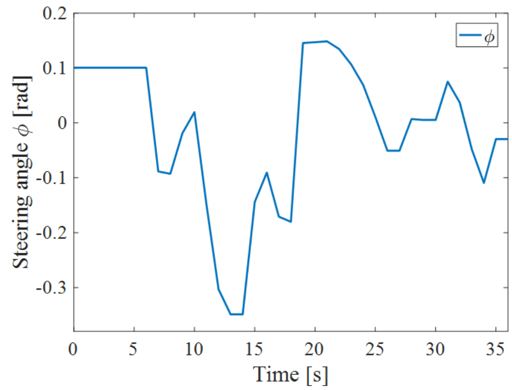

Fig. 12. Amount of the steeing angle value

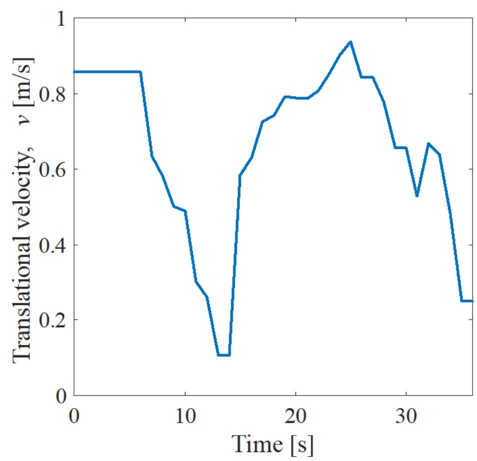

Fig. 13. Amount of the robot velocity

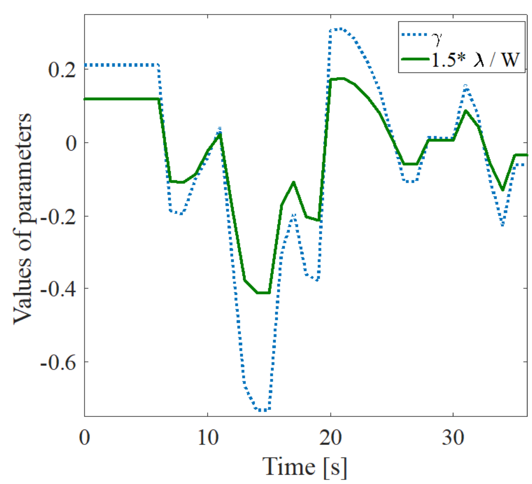

Fig. 14. Amount of parameters, γ and λ, of human tracking