Indonesian Journal of Electrical Engineering and Computer Science
Vol. 26, No. 3, June 2022, pp. 1338~1350
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v26.i3.pp1338-1350

Design of elderly-assistant mobile servant robot

Minh Son Nguyen, The Tung Than, Tri Nhut Do, Hoai Nhan Nguyen
Faculty of Computer Engineering, University of Information Technology
Vietnam National University Ho Chi Minh City (VNUHCM), Ho Chi Minh City, Vietnam

Article Info

Article history:

Received Sep 20, 2021
Revised Mar 28, 2022
Accepted Apr 3, 2022

Keywords:

Embedded system
Face recognition
Mobile robots
Real-time image processing
Tracking elderly

ABSTRACT

Recently, elderly population increasing worldwide has put higher pressure on health-care providers and their families. The advent of elderly care robots will reduce that pressure. In this paper, a design of mobile servant robot with integrated tracking algorithm in order to assist the elderly by companionship is proposed not only to help families take care of their elderly at home but also reduce the pressure on health-care providers. The proposed robot is based on humanoid structure and AI-embedded-GPU controller. The design allows the robot to follow the elderly and accompany them in real-time. In addition, the video streaming algorithm with the pipeline mechanism is integrated on robot controllers so that the owner interacts with the elderly through the internet. The robot controller is embedded into hardware of 128 graphics processing unit cores and 4 ARM Cortex-A9 cores in order to execute convolutional neural network (NCNN) algorithms for elderly recognition and body tracking. The processing speed at 14 fps of video stream in real-time. The proposed robot can move on uneven surfaces with a speed at 0.21 m/s and an accuracy over 90%. However, the video stream processing speed is able to be reduced at 15 fps and latency less than 415 ms when four users appear concurrently.

Corresponding Author:

Minh Son Nguyen
Faculty of Computer Engineering
University of Information Technology at VNUHCM
KM20 Quarter 6, Linh Trung Ward, Thủ Đức City, Ho Chi Minh City, Vietnam
Email: sonnm@uit.edu.vn

1. INTRODUCTION

According to the World Population Prospects 2019 [1], the world population is aging with the group of the people who aged more than 65 years increasing the fastest. It is forecasted that by 2050, 1/6 of the world's population will be over 65 years old (about 16%) and a quarter of the population living in Europe and North America will probably be over 65 years old, which means there will be a person aged more than 65 for each group of four. Also according to this report, in 2018, for the first time in human history, people aged over 65 outnumbered children under 5 in the world. Elderly people aged 80+ are expected to nearly triple from 143 million in 2019 to 426 million in 2050.

The aging population is increasing world-wide which renders their needs an important matter to health providing authorities, agents of governments, caregivers, and families. This leads to the emergence of healthcare robots which have a role in assisting older adults to complete their daily activities, helping to monitor behavior and health of the elderly, and acting as a companion when they are alone [2]. In the near future, the world will have a serious shortage of aged care workers, which will cause the cost of elderly care to rise, creating a burden for families and carers. Therefore, elder care robots (ECRs) is an adequate compensation for that shortage, it will replace carers, help and supervision of the elderly (WHO, 2016).

An elder care robot needs such attributes as functions that are well designed to satisfy the requirement of operation, emotion and social needs. These features of the assistive robot system influence elder perceptions and attitudes. There is some research focused on solving robot design for the purpose of assisting the elders [3]-[10]. The medical robotics help the elderly in two ways [3]: i) Service type robot supports independent living such as eating, bathing, toileting and getting dressed, and mobility, providing household maintenance, monitoring of those who need continuous attention and maintaining safety; and ii) Companion type robots enhance health and psychological wellbeing of elderly users by providing companionship. Some versions of companion robots which were made by different companies such as Aibo, Paro, iCat, Pearl, nursebot, Care-o-bot, Homie, Huggable, and Robocare. They can be programmed to be in many ways with elderly, assistive robotics, healthcare or health and care. Pepper [4] is a humanoid robot, with 17 joints (the elbow, shoulder, and hip) for being capable of exhibiting body language. To move around smoothly, it is equipped with three omnidirectional wheels. It is also capable of perceiving and interacting with its surroundings, moving around, analyzing people’s expressions and voice tones of humans around it. Therefore, Pepper is suitable for being an elder care. It also has a tablet on the chest, and many sensors: 2D cameras, 3D sensor, lasers, sonars, infra-red sensors, etc. Pepper is able to sense basic environment status, for instance, human’s presence or simple objects thanks to built-in system commands. Until now, pepper model provides a field of service mainly based on a dialogue system. To apply the robot model in caring for elders, it is necessary to add more functions to taking care of elders in practice. Moreover, Pepper's conversation topic was restricted because of the limitation of the dialogue system. So, it is needed to build up the database to enlarge the topic.

In the papers related to robot care and elderly of Masaki Onishi, a research team from Riken Research Institute (Japan) [5]: the companion human-interactive robot (RI-MAN) serving humans can move, track and perform some functions such as carrying people, listening to the rhythm breathing and distinguishing some smells. As stated by the research group on smart robots Gregoire Milliez in the US [6], multi-function smart healthcare robots used in the home often have the function of supporting monitoring, detecting abnormal situations, controlling home devices, monitor elderly care, remind posture and schedule, and provide multimedia social interaction in the network environment. Artificial intelligence roBOt (AIBO) is designed as a mobile and autonomous robot for an entertainment purpose (made by Sony) [7]. The robot behavior is programmable; it has a hard plastic exterior, equipped with many sensors such as a camera, touch sensors, infrared and stereo sound. Robot gestures are performed by actuators such as four legs, a moveable tail, and a moveable head. AIBO is programmed to play and interact with human beings. It was used in studies with the elderly in order to try to assess the effects on the quality of life and symptoms of stress [7]. A baby harp seal robot (PARO) is designated as a companion robot [8]-[10]. This soft seal robot is targeted at the elderly. The robot behavior is programmable, and it also includes a touch sensor, an infrared sensor, stereoscopic vision and hearing. Robot actuators include eyelids, upper body motors, front paw and hind limb motors. Pearl is a mobile robot [11] named as nursebots (developed by Carnegie Mellon University). It helps the elderly to navigate through the nursing facility. It has a user-friendly interface, providing advice and cognitive support for the elderly. Other elder care robots such as CareO-Bot [12], RoboCare [13], and HUGGABLE robots [14] are also listed as companion robots.

This paper proposes a design of a robot to accompany the elderly. The tracking algorithm is integrated in the robot controller to follow the elderly. By this solution, the designed robot helps families take care of their elderly at home but also reduce the pressure on health-care providers. Simultaneously, the robot performs both processes of accompanying the elderly and streaming the elderly's image data over the internet. The elder assist robot determines which objects need to be monitored. To increase the ability to distinguish between different objects, we used faces, this is the feature on each person that will have the most obvious difference. After identifying the target object, the system will conduct the process of tracking the object on the image frame, calculating the parameters from which to control the robot to follow the elderly object (robot maintains a distance 2 m between it and the object).

The main contributions of this paper are: i) An embedded system integrated identification algorithm on GPU so that this robot can move to the elderly and accompany them in real-time (≈14 fps), ii) Mobile Servant Robot which can follow the elderly when multiple objects appear in view of the camera, and iii) the video streaming algorithm with pipeline mechanism is integrated so that the owner can interact with their elderly through the internet.

The remainder of the paper is structured as follows: section 2 describes the entire design of the robot including hardware and application software; algorithms; section 3 shows the experimental scenarios in detail, and collects the robot's corresponding results with the experimental scenarios and statistics; the evaluation of experimental results is presented in section 4 and finally the conclusion on the topic proposed in this paper is summarized in section 5.

2. RESEARCH METHOD

2.1. System overview

Based on some research on object tracking robots such as Masaki Onishi team's companion robot [5], Companion Robot [6], and OpenBot [15]. We designed a robot capable of tracking elderly when they are alone at home with the ability to move on flat surfaces at a speed equal to that of elderly people. The robot uses input image processing on a single camera to identify and confirm the object, and then conducts the process of tracking and accompanying the object. At the same time, the robot also transmits image data to their relatives via the internet. These properties are referenced from the studies in the paper [16].

2.2. Robot hardware architecture

2.2.1. Modeling of the mobile server robot

The structure of the Mobile Servant Robot accompanying the elderly is designed with two parts: the body like the Humanoid Robot model and the moving part is designed according to the crawler mechanism as shown in Figure 1 [17]. This crawler mobile robot structure is widely used due to its simplicity, flexible movement, convenient to control and driven by a DC electric motor. This crawler has an additional function of keeping the robot balanced when moving. With the purpose of convenience for indoor movement and excellent balance, we realize that using the crawler part instead of the legs will help the robot be more agile and meet the requirements set forth.

Robot kinematic diagram is illustrated in Figure 2. Robot has 2 arms, left one and right one. Each arm has 3 dof (degree of freedom). Robots can move based on tracked skid locomotion. The form of skid steering (like the army tank operation) is used to control the robot orientation. By controlling spinning wheels (connected to left, right tracks) at different speeds the robot can be turned left or right. In the case of turning left, the right track rotates at higher speed compared to the left one, and inversely, in the case of turning right, the left track rotates at a higher speed than the right one. Skid steering has the properties of high maneuverability, a simple, robust structure. Therefore, the proposed robot has good mobility on a lot of terrains.

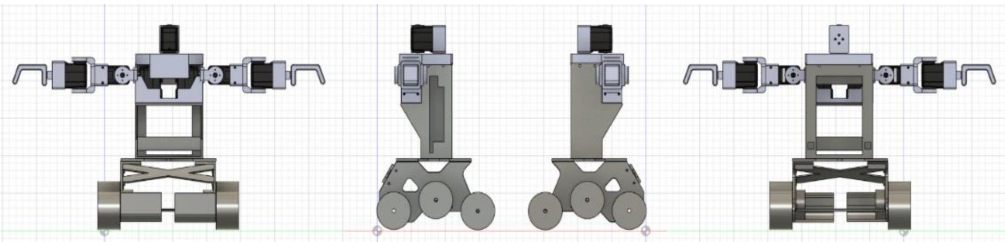

Figure 1. 2D model of robot in this study

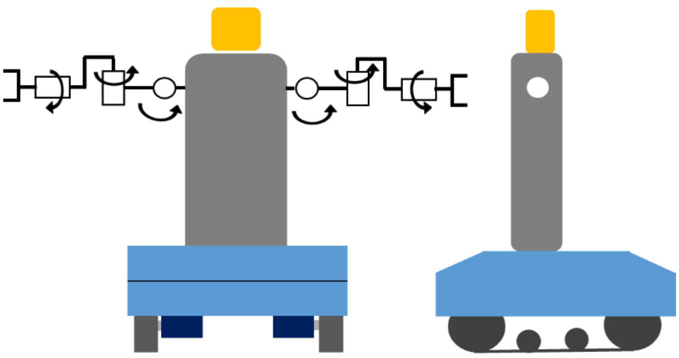

Figure 2. Robot kinematic diagram

The model of mobile servant robot in this research is subscribed simply as a moving wheeled object (Figure 3), operating on a flat horizontal surface. The total degrees of freedom is 3 including 2 degrees of freedom representing the robot position in the plane and a degree of freedom is a rotation of the robot around the vertical axis which is perpendicular to the flat horizontal surface.

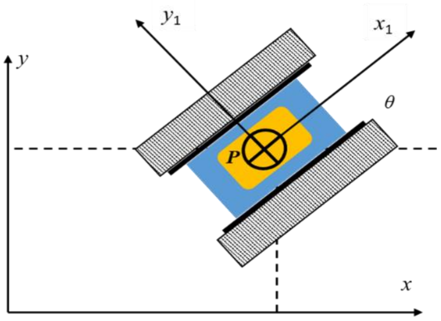

Figure 3. Skid-steering robot locomotion

By assigning the reference coordinate frame (x, y), can be considered as a global frame, and robot body frame (O1, x1, y1), as shown in Figure 3. Robot configuration (position of a point P on robot and robot orientation) can be defined by the relationship of these coordinate frames. Position of the robot is defined by point P which is the origin of the frame (O1, x1, y1); orientation of the robot can be defined by angle θ. A column vector ξ of which elements are x, y, θ is formed to describe robot configuration.

$$\xi = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \quad (1)$$

A relationship of frame (O,x,y) and frame (O1, x1, y1) is described by (2).

$$R(\theta) = \begin{bmatrix} cos\,\theta & sin\,\theta & 0 \\ -\sin\theta & cos\,\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

The principle of differential transmission of two active wheels: The model describing the binding conditions of this differential drive requires the parameters (distance L between two wheels, radius r of each wheel). The angular velocity vector of the two wheels corresponds to the two components of the following vector $u = (u_{right}, u_{left})$; These 2 components determine the angular velocity of the left and right wheels (unit: radians/s). Consider the moving cases of the mobile robot (Figure 4(a) and Figure 4(b)):

- If $u_r = u_l > 0$, the robot moves straight forward.
- If $u_r = -u_l \neq 0$, the robot rotates in a clockwise direction thanks to two active wheels that rotate in the opposite direction.

Figure 4. Mobile robot movement: (a) the robot moves in a straight line, 2 wheels rotate in the same direction and (b) rotating motion in place, 2 wheels rotate in opposite directions

The (3) describes the velocity conversion along the x, y axes and the rotational speed around the vertical axis The angular speed θ is proportional to the change in the angular speed value of the two wheels. The rotation speed of the robot is directly proportional to the wheel radius and inversely proportional to the distance between the two wheels.

$$\dot{x} = \frac{r}{2}(u_l + u_r)\cos\theta$$
$$\dot{y} = \frac{r}{2}(u_l + u_r)\sin\theta$$
$$\dot{\theta} = \frac{r}{L}(u_l - u_r) \quad (3)$$

2.2.2. Hardware control system for mobile servant robot

The hardware system is designed based on the principle of the embedded system (see Figure 5), in which the central controller has 2 processors including: 1 ARM Cortex A9 processor for motor control and peripheral communication. [WIFI, collect images from Camera sensor…] and the GPU processor executes convolutional neural network algorithms with the image data. In addition, the hardware system has a built-in 720p HD camera sensor that collects input images for the robot to process and transmits control signals to the controller board, then controls the motor to move. In addition, the system also has a Wifi-Adapter attached to facilitate debugging, control the Robot in manual mode via a computer and perform the function of streaming video over the internet.

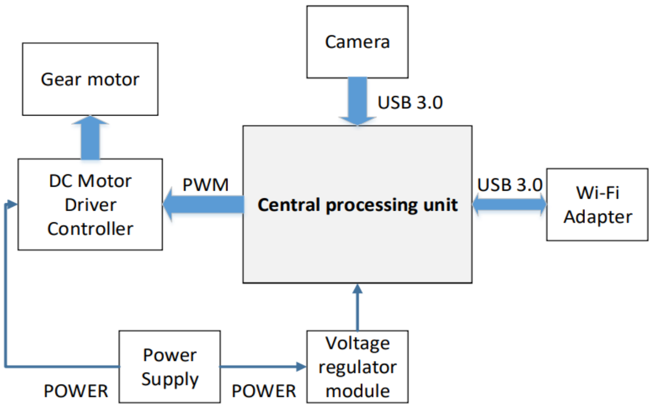

Figure 5. Hardware system design for robot controller

2.3. Software system

2.3.1. Software system overview

After referring to two types of robots with human tracking functions, we found that Openbot [15] has fast moving speed along with real-time autonomous navigation, but still cannot distinguish objects and monitor when multiple objects appear. As for the Miura group's robot [18], they use the SIFT feature to identify and track objects, this depends quite a lot on the clothes which the elderly wear at that time, difficult to recognize the original object along with the processing speed of the system is also not high. So we used the face recognition algorithm because the face is the most distinctive feature of each person, it will increase the ability to identify the object. Moreover, we will use the tracking with correlation filters (KCF) [19] tracker, this is a lightweight algorithm, capable of tracking in environments with many objects, without using too many hardware resources, resulting in achieving real-time speed. In addition, an advantage of this system is the ability to stream video over the internet, which allows the relatives to monitor the activities of the elderly at home, the WebRTC [20] algorithm is used due to its high security, is supported across laptop, PC or mobile devices, and WebRTC does not need supporting applications or Plugins.

The main function of the software system is to monitor and accompany the elderly, but first the robot needs to determine which objects need to be monitored. To increase the ability to distinguish between different objects, we used faces, this is the feature on each person that will have the most obvious difference. After identifying the target object, the system will conduct the process of tracking the object on the image frame, calculating the parameters from which to control the robot to follow the elderly object (distance between Robot and object will be maintained within 2 m) [16]. Simultaneously with the process of accompanying the elderly, the system will stream the elderly's image data over the internet. Figure 6 shows the overall control software algorithm for the companion robot in this topic. In this algorithm, the robot's software system performs the following steps:

- S1: Collect image data from the camera. This step uses a parallel execution algorithm aimed to stream video over the internet and perform object tracking. The technique of using a virtual video stream shared from the physical video stream allows multiple processes to execute video processing at the same time and still achieve real time speed.
- S2: Identify the object to be tracked. This step uses a face recognition algorithm that uses a convolutional neural network for identification.

– S3: Object tracking. The tracking algorithm to track and control the moving robot to accompany the object is integrated into the controller. In this algorithm, if the identified object is lost, the algorithm will return to the facial recognition step in S2.

In order for the processing system to achieve real-time capabilities, we implement a convolutional neural network algorithm that performs optimally with GPUs according to NCNN technology. The technology is a high-performance neural network computation and inference framework optimized for mobile and embedded platforms. Besides, NCNN technology does not depend on third parties, so it can execute cross-platform on mobile devices and embedded computers [21].

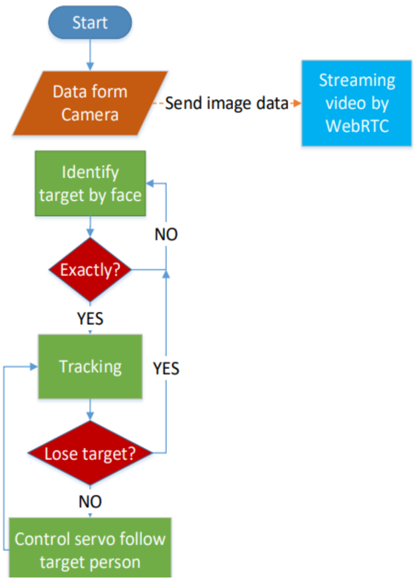

Figure 6. Software system processing flowchart

2.3.2. Algorithms used in software system

Figure 7 shows three algorithmic schemas of three algorithms designed to be integrated on the companion moving robot for this study, which include: video streaming algorithm using WebRTC technology, face recognition algorithm using RetinFace [22] and MobileFaceNet [23] and SSD_MobileNet [24] user body detection algorithm. In the following, the research team will present the processing steps of the software system.

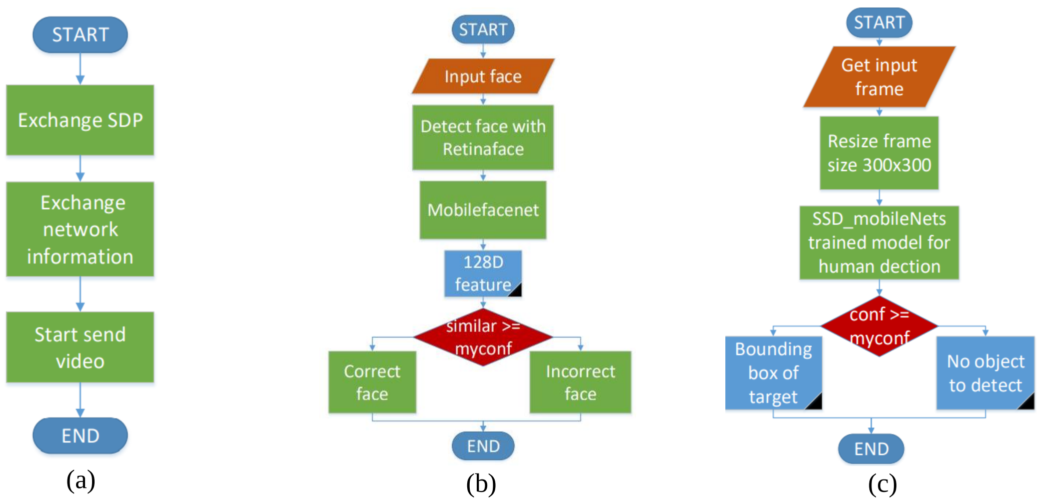

Figure 7. Software system processing flowchart of (a) Streaming video algorithm, (b) Facial recognition algorithm, and (c) Human detection algorithm

Firstly, with a view to performing a video streaming function that can meet real-time and low-latency requirements, the WEB Socket technique with a pipeline mechanism is implemented for fast video processing over the internet infrastructure. In addition, the technique of using a virtual video stream on shared memory from the physical video stream allows multiple processes (clients) to execute video processing at the same time and meet the real time speed (Figure 7(a)).

Secondly, the face recognition algorithm implemented on GPU cores, the RetinaFace model, is applied to search for faces appearing in the frame and then extract the object's face as input for face verification. The faces are recognized with the MobileFaceNet algorithm through the RetinaFace model to enhance the accuracy of the recognition algorithm. The output of the MobileFaceNet model will be a vector with 128 feature values of that face. After having the feature vector including the features of the face, compare it with the model face features of the object to be recognized. Depending on the similarity of the object's face, whether it is the face that needs to be recognized or not (see Figure 7(b)).

Then, from output of the face and body recognition algorithm, the Robot compares the coordinates of the face with the coordinates of the body, the body closest to the face will be the coordinates of the body to be tracked in the current frame. As the tracking algorithm KCF–DSST is not able to resize the bounding box of object very well, we integrate this algorithm with SSD_MobileNet algorithm that pre-trained with a training dataset taken from Caltech Pedestrian (Figure 7(c)) in order to resolve the problem of the tracking algorithm, as shown in Figure 8. With an aim of increasing the accuracy of the object tracking algorithm, determining the distance of the object in the image frame and the direction of movement is very important. The algorithm determines the distance according to (4) and as shown in Figure 9.

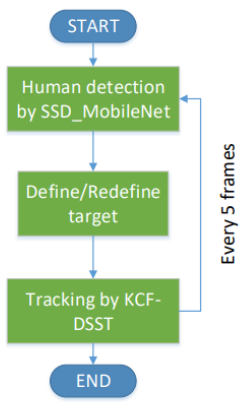

Figure 8. The flowchart algorithm of object tracking according to KCF–DSST [25]

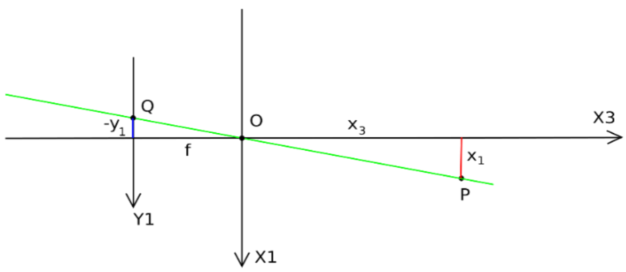

Figure 9. Pinhole camera model [26]

In which, measuring the distance using the pinhole camera model formula to calculate the distance from the camera to the object, because this is a simple formula, easy to apply, can be applied with one camera. From there, the formula for calculating the distance to the object in the frame is determined as (4):

$$\frac{x}{f} = \frac{X}{d} \quad (4)$$

where:
d: distance to object (mm)
x: size of object in frame (mm)
X: size of actual object (mm)
f: camera focal length (mm)

Finally, in order to determine the moving direction of the object in the frame, the robot needs to integrate an algorithm to determine the direction of the object's movement so that it does not lose track when following. In the algorithm to determine the direction of moving left or right, the algorithm compares from the center position of the frame to the center position of the object, this is executed for each frame on GPU processors. At the same time, to determine whether the object is moving backward or forward, we use the algorithm to determine the object distance according to Figure 9 and (4). From the coordinates of the object extracted from the above steps, the robot determines the distance between the center of the object and the center of the frame, determines the direction of movement, thereby controlling the motors to follow the object (Figure 10). In this study, we use a distance of 2 m as the minimum distance between the robot and the object because this is the ideal distance for high-accuracy face recognition algorithms as described in the previous section. In fact, because using a conventional camera it is difficult to determine the exact distance to the moving object, so we assume in the algorithm for the minimum distance between the robot and the tracked object in the 1.8-2.5 m range, this is a method to minimize errors when controlling the Robot.

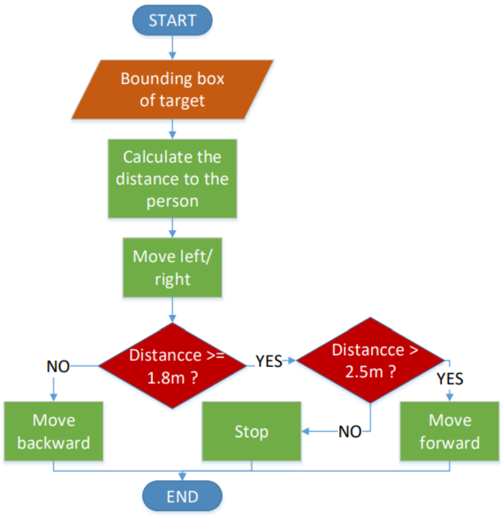

Figure 10. The flowchart algorithm of controlling target tracking robot

3. EXPERIMENTS AND RESULTS

In order to prove the correctness of the proposed functions in section 2, we proceed to build test scenarios and evaluate the robot with 4 phases. Firstly, we determine the robot's ability to distinguish objects by evaluating accuracy and execution speed of the facial recognition algorithm (section 3.1). Secondly, the tracking capacity of the robot in the frame sequence is assessed through accuracy and execution speed of moving person detection algorithm, moving object tracking (section 3.2, 3.3). In addition, speed and latency of internet video streaming algorithms with common WEB browser applications in section 3.4. Finally, Section 3.5 demonstrates the capacity of the robot in a real-world environment based on evaluations of the processing rate of one frame, moving speed and movement accuracy of the robot companion with elderly objects. Following are the experimental scenarios and the obtained results.

3.1. Experiment with facial recognition algorithm

3.1.1. Experimental scenario

Face data is trained on a computer with a powerful GPU performance, then the facial feature data is integrated on the Robot's Controller as the basis for the facial recognition algorithm to check for duplication in the image frame taken directly from the Camera sensor. In this experiment, we let the robot learn the face and body of a man, 175 cm tall, weighs 60 kg, and wears glasses. We experiment 1000 times on different environments: i) faces at different angles with increasing distances from 0.5 to 2.5 m., ii) ambient light above 100 lux, and iii) 2, 3, 5 faces appear in the same frame.

3.1.2. Experimental results

From Table 1, we realize that the algorithm will work most effectively with the conditions as: the object has a distance in the 0-2.2 , range to the robot; when the robot detects the object's face is larger than ¾ standard face and the surroundings has good brightness (over 100 lux).

Table 1. Facial recognition experiments with different distances and angles

Distance	Opposite camera	Turn face 0 to 45 degrees left	Turn face 0 to 45 degrees right	Face up and down
0.5 m	100%	81.50%	83.50%	83.50%
1 m	100%	76.10%	78.60%	72%
2 m	99.90%	58.00%	58.40%	70%
2.5 m	47.40%	17.50%	18.20%	29.70%

As reported by Table 2, the processing time of the algorithm will increase as the number of faces increases. However, for the application of tracking one single object, this robot is capable of recognizing faces at a speed of 25 fps with an accuracy of 100% in the frontal direction. However, Robot can recognize more faces in the same context with slower speed [less than 20 fps].

Table 2. Experiment in the environment with many faces facing the robot with a distance of 1-2 m

Number of Faces in front of Robot	1	2	3	5
Accuracy	100%	100%	89.2%	81.6%
Speed	39 ms	49 ms	54 ms	82 ms

3.2. Experiment with moving human detection algorithm

3.2.1. Experimental scenario

Similar to the face recognition algorithm. In this experiment, we use a human body dataset and perform training on a GPU-enabled computer. The model is then stored in the memory of the Robot controller. Carry out the experiment 1000 times on a man with the same parameters as experiment A according to the scenario: change the standing-sitting postures with different angles, rotate the whole body 360 degrees and move the object.

3.2.2. Experimental results

According to the experimental results of Table 3, the robot accurately detects objects in standing-sitting positions over 91.26% and the processing speed is 45 ms (corresponding to 22 fps). When the object moves, the robot's ability to detect will be reduced to 80% because the frame is blurred when the object moves (human detection algorithm cannot detect it correctly).

Table 3. Experimental results of detecting people in a standing-sitting position within a distance of 2 m from the robot

Experimental case	Standing postures	Sitting postures	Moving object	360 degrees
Accuracy	91.26%	100%	80%	91.6%
Speed	49 ms	44 ms	47 ms	43 ms

3.3. Experiment with moving object algorithm

3.3.1. Experimental scenario

To evaluate the robot's moving object tracking algorithm. We offer an environment scenario where one person needs to be monitored, an environment with many people including one person who needs to be monitored and accompanied. This experiment shows the ability to track objects in each image frame.

3.3.2. Experimental results

After experimenting 100 times with each case (Table 4), we draw the following conclusions: the average speed of the algorithm is 27.5 ms (corresponding to 36 fps), the processing time in the 13-62 ms range, depending on the size parameter of the object to be tracked. The Table 4 shows the feasibility of the tracking algorithm.

Table 4. Experimental results of moving object tracking algorithm

Parameter Experimental scenario	Bounding box size (px×px)	Processing time (ms)
One object is an elderly who can move on his/her own	311×136	20
One object is an elderly, using a walker	94×109	15
Two objects including the elderly	181×424	62
More than 2 objects	34×74	13

3.4. Experiment with streaming video algorithm

3.4.1. Experimental scenarios

Experiments with the robot's video streaming capability so that relatives of the elderly can watch simultaneously. Experiment with the number of relatives connecting to the Robot simultaneously from 1 to 4 people on popular web browsing applications such as: Microsoft Edge, Firefox, Coc Coc, Safari, and Google Chrome. The robot is connected to the IEEE 802.11 b/g/n wireless WIFI network with a 20 Mbps internet packet. With the experiments below, we all use a frame resolution of 640×480 to transmit between the robot and the elderly's relatives on the internet infrastructure.

3.4.2. Experimental results

Video streaming algorithm using WebRTC gives impressive results at 15 fps, the average streaming latency when used by a user is 300 ms. Can allow four users to use at the same time with latency less than 415 ms and is supported on almost all popular web browsers. It means that the four people in the family are able to view the real time video from mobile server robot [checking the situation of elders in home] at the same time. The Table 5 shows the feasibility of the tracking algorithm.

Table 5. Stream delay on web browsers when increasing the number of users (unit: ms)

Number of users Web browser	1	2	4
Microsoft Edge	305	320	354
Coc Coc [VN WEB Browser]	270	290	346
Firefox	326	370	415
Google Chrome	300	313	366
Safari	286.67	318	350

3.5. Experiment with moving accuracy of robot companion with elderly

3.5.1. Experimental scenario

To evaluate the movement accuracy of the companion robot, the following scenarios were tested by us on a male, 175 cm tall, weighs 60 kg, wearing glasses in 3 cases:

– Case 1: Mobile Servant Robot operates in space with only target objects.
– Case 2: Robot tracks moving object.
– Case 3: Robot operating in space with many people.

The scenario on Figure 11 is built to evaluate the robot's tracking ability in an environment with many objects. The target object will move in order from point 0 to point 5 (green) corresponding to which the robot will also move to the corresponding points (orange). The initial distance between two corresponding points is 2 m. As a result, the Robot always moves and follows the object using the KCF-DSST algorithm with an accuracy of over 90%.

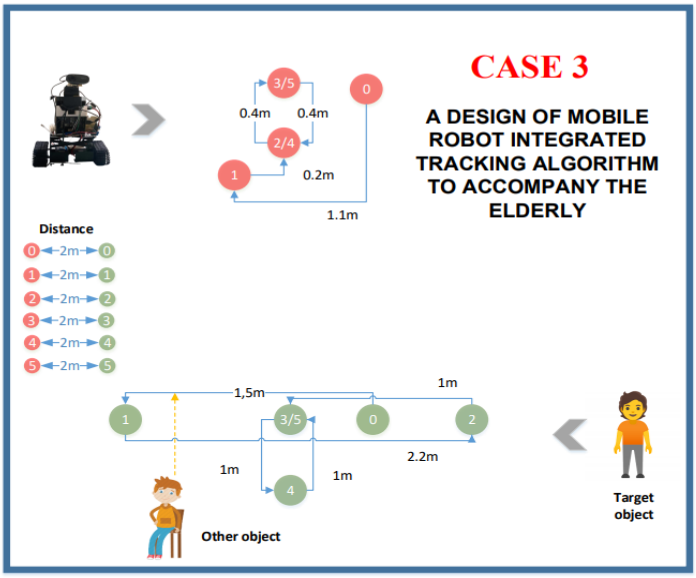

Figure 11. Case 3 of the experiment, in space there are many people. The image depicts the direction of the robot's movement, corresponding to the direction of the object's movement

3.5.2. Experimental results

The results obtained from the experimental scenario Robot companion with the elderly and the parameters (Table 6): the average processing speed of scenarios per frame is 70 ms (corresponding to 14 fps).

Moreover, the average distance to the objects of the scenarios is 2.17 m. Cases 1 and 3 are scenarios built to test the ability to track objects moving sideways, the speed at which the Robot can track and follow without losing the object is 0.33 m/s. Meanwhile, Case 2 is a scenario built to test the robot's ability to follow when people travel long distances with the robot's moving speed of 0.21 m/s.

Table 6. Results obtained from experimental scenarios robot accompanies the elderly

Parameter Case	Processing rate of one frame (ms)	Distance to object (m)
1	66	2.18
2	66	2.2
3	78	2.13

4. EVALUATION

In this study, we have mentioned 2 types of mobile robots of 2 groups Matthias Müller [15] and Miura [18] that are applied in the house that have object tracking functions. According to the Table 7, our mobile server robot has the superiority in processing speed based on the number of frames per second compared to the other 2 robots, as well as the robot's average moving speed is close to the moving speed of the elderly for continuous monitoring and follow-up. We also developed the ability to recognize objects by the unique feature of each person's face to increase the ability to distinguish objects in reality. In addition, the function of streaming videos over the internet to share with relatives is also an advantage of this robot. However, in this study, the team needs to increase the robot's average moving speed and add an obstacle avoidance algorithm. Besides, the robot in this study has a number of other advantages as follows: capable of operating in an environment where many people appear together, the latency of video streaming algorithms is relatively low (<415 ms) with frames 640×480 images can be used for 4 relatives in real time.

Table 7. Robot comparison of this study with some other companion robots

Type of Robot Specifications	Our robot	OpenBot [15]	Robots of the Miura group [18]
Compute	ARM Cortex-A9	Smartphone	Core2Duo
Camera	Conventional camera	Smartphone camera	Stereo camera
Average speed (m/s)	0.21	1.5	0.3
Tracking a moving object	Yes	Yes	Yes
Frames per second (fps)	13-20	>10	<10
Monitor when multiple objects appear	Yes	No	Yes
Ability to distinguish tracking objects	Yes	No	Yes
Streaming video over the internet: - WebRTC technology - Latency: 415 ms with 640×480 resolution	Yes	No	No
Obstacle avoidance	No	No	No
Price	Medium (260$)	Lower	Higher
Weight (kg)	1.51	0.7	N/A
Working time with 1800 mAh Battery (min)	83	45	N/A
Dimensions (length x width x height) in cm	24×18×29.6	24×15×12	N/A

5. CONCLUSION

In this paper, a research on designing hardware and software algorithms for the Mobile Servant Robot to accompany the elderly to support relatives monitoring as well as interacting with the elderly at home. This robot has a body designed like a humanoid robot and a lower body with a crawler mechanism to be able to move on terrain with uneven surfaces. The Robot's Controller is designed based on the Embedded System integrated identification algorithm on the GPU so that this Robot can move to the elderly and accompany them in real-time. Besides, the robot has a function that allows relatives to view images of the elderly at home through the video streaming function according to the pipeline mechanism. Through experimental scenarios, the robot has the function of recognizing elderly people through facial recognition and conducting object tracking to achieve real time speed up to 14 fps using a convolutional neural network (NCNN). In addition, the robot also transmits image data of the elderly to their relatives via the internet at a speed of 15 fps, transmission delay of less than 415 ms for 4 simultaneous accesses.

ACKNOWLEDGEMENTS

This research was supported by The VNUHCM-University of Information Technology's Scientific Research Support Fund, under project grant number D1-2022-04.

REFERENCES

[1] The Department of Economic and Social Affairs of the United Nations Secretariat, "World Population Prospects," 2019. [Online]. Available: https://population.un.org/wpp/, accessed: 2020-09-01.

[2] S. H. Hosseini and K. M. Goher, "Personal care robots for older adults: An overview," *Asian Social Science*, vol. 13, no. 1, 2017, doi: 10.5539/ass.v13n1p11.

[3] K. M. Goher, N. Mansouri, and S. O. Fadlallah, "Assessment of personal care and medical robots from older adults' perspective," *Robotics and biomimetics*, vol. 4, no. 5, 2017, doi: 10.1186/s40638-017-0061-7.

[4] A. K. Pandey and R. Gelin, "A mass-produced sociable humanoid robot: pepper: the first machine of its kind," in *IEEE Robotics & Automation Magazine*, vol. 25, no. 3, pp. 40-48, Sept. 2018, doi: 10.1109/MRA.2018.2833157.

[5] M. Onishi, Z. Luo, T. Odashima, S. Hirano, K. Tahara and T. Mukai, "Generation of human care behaviors by human-interactive robot RI-MAN," *Proceedings 2007 IEEE International Conference on Robotics and Automation*, 2007, pp. 3128-3129, doi: 10.1109/ROBOT.2007.363950.

[6] G. Milliez, "Buddy: A companion robot for the whole family," *HRI '18: Companion of the 2018 ACM*, 2018, doi: 10.1145/3173386.3177839.

[7] M. Fujita, "AIBO: Toward the era of digital creatures," *The International Journal of Robotics Research*, vol. 20, no. 10, pp. 781-794, 2001, doi: 10.1177/02783640122068092.

[8] K. Wada and T. Shibata, "Living with seal robots-its sociopsychological and physiological influences on the elderly at a care house," in *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 972-980, Oct. 2007, doi: 10.1109/TRO.2007.906261.

[9] K. Wada, T. Shibata, T. Saito, and K. Tanie, "Effects of robot assisted activity to elderly people who stay at a health service facility for the aged," *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)*, 2003, pp. 2847-2852 vol.3, doi: 10.1109/IROS.2003.1249302.

[10] M. Pollack *et al.*, "Pearl: A mobile robotic assistant for the elderly," *AAAI Workshop on Automation as Eldercare*, 2002.

[11] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, "Towards robotic assistants in nursing homes: Challenges and results," *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 271-281, 2003, doi: 10.1016/S0921-8890(02)00381-0.

[12] B. Graf, M. Hans, and R. D. Schraft, "Care-O-botII—Development of a Next Generation Robotic Home Assistant," *Autonomous Robots*, vol. 16, no. 2, pp. 193-205, 2004, doi: 10.1023/B:AURO.0000016865.35796.e9.

[13] S. Bahadori *et al.*, "RoboCare: an integrated robotic system for the domestic care of the elderly," *Proceedings of workshop on Ambient Intelligence AI, Pisa, Italy*, 2003.

[14] W. D. Stiehl *et al.*, "The huggable: a therapeutic robotic companion for relational, affective touch," *2006 3rd IEEE Consumer Communications and Networking Conference*, 2006, pp. 1290-1291, doi: 10.1109/CCNC.2006.1593253.

[15] M. Müller and V. Koltun, "OpenBot: turning smartphones into robot," *Computer Vision and Pattern Recognition*, 2020, doi: 10.48550/arXiv.2008.10631.

[16] M. N. A. Bakar and M. F. M. Amran, "A study on techniques of person following robot," *International Journal of Computer Applications*, vol. 125, no. 13, pp. 27-30, Sep. 2015, doi: 10.5120/ijca2015906165.

[17] Van Tan Nhat Vo, Tan Tan Nguyen, "Motion Generation for Humanoid Based on Human Gesture using Kinect," *Vietnam Mechanical Engineering Journal*, 2017, ISSN 0866- 7056, pp. 79-87.

[18] J. Satake, M. Chiba, and J. Miura, "A SIFT-Based person identification using a distance-dependent appearance model for a person following robot," *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2012, pp. 962-967, doi: 10.1109/ROBIO.2012.6491093.

[19] J. F. Henriques, R. Caseiro, P. Martins and J. Batista, "High-speed tracking with kernelized correlation filters," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 583-596, 1 March 2015, doi: 10.1109/TPAMI.2014.2345390.

[20] J. Cui and Z. Lin, "Research and implementation of WebRTC signaling via websocket-based for real-time multimedia communications," in *5th International Conference on Computer Sciences and Automation Engineering (ICCSAE 2015)*, 2015, pp. 374-380, doi: 10.2991/iccsae-15.2016.72.

[21] S. Lin, R. Ji, C. Chen, and F. Huang, "ESPACE: accelerating convolutional neural networks via eliminating spatial and channel redundancy," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, pp. 1424-1430, 2017.

[22] J. Deng, J. Guo, E. Ververas, I. Kotsia and S. Zafeiriou, "RetinaFace: Single-shot multi-level face localisation in the wild," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 5202-5211, doi: 10.1109/CVPR42600.2020.00525.

[23] S. Chen, Y. Liu, X. Gao, and Z. Han, "MobileFaceNets: Efficient CNNs for accurate realtime face verification on mobile devices," *Chinese Conference on Biometric Recognition*, 2018, pp. 428-438 doi: 10.1007/978-3-319-97909-0_46.

[24] S. Arabi, A. K. Haghighat and A. Sharma, "A deep learning based solution for construction equipment detection: from development to deployment", *Computer Vision and Pattern Recognition*, 2019, doi: 10.48550/arXiv.1904.09021.

[25] M. Danelljan, G. Häger, F. Shahbaz Khan, and M. Felsberg, "Accurate scale estimation for robust visual tracking," in *Proceedings of the British Machine Vision Conference (BMVC)*, 2014, doi: 10.5244/C.28.65.

[26] K. Hata and S. Savarese, "CS231A Course Notes 1: Camera Models," *CS231A: Computer Vision, From 3D Reconstruction to Recognition*, 2021.

BIOGRAPHIES OF AUTHORS

The-Tung Than graduated from The University of Information Technology in 2021 with a major in Embedded System and Robot, achieved a GPA of 8.34/10. Currently, Mr. THAN is a Teaching Assistant in the Faculty of Computer Engineering at University of Information Technology-VietNam National University at HCMCity. He is a careful and hard-working person. He is always willing to learn new things and happy to create something useful for everyday life. His research is focussed on fields of Embedded System Design, AI for IoT, Autonomous Mobile Robots. He can be contacted at email: tungtt@uit.edu.vn.

Tri-Nhut Do received the B.Eng. degree in electrical electronics engineering and the M.Eng. degree in technology cybernetics engineering from the Ho Chi Minh City University of Technology, Vietnam, in 2002 and 2005, respectively, and the Ph.D. degree in electrical engineering from the University of Ulsan, South Korea, in 2012. From 2013 to 2014, he was a Post-Doctoral Fellow under the supervision of Prof. R. Phan with the Security Laboratory, Multimedia University, Malaysia. From 2015 to 2017, he was a Post-Doctoral Fellow under the supervision of Prof. U-X. Tan and Prof. C. Yuen with the Robotics Innovation Laboratory, Pillar of Engineering Product Development, Singapore University of Technology and Design, Singapore. From 2018 to 2020, he was a lecturer with the Faculty of Technology and Engineering, Thu Dau Mot University, Thu Dau Mot City Vietnam. Since 2021, he has been a lecturer with the Faculty of Computer Engineering, University of Information Technology, HCMC Vietnam. His research interests include indoor localization, human daily activities tracking, robotics, sensor fusion, and human emotion recognition for security. He is a recipient of the outstanding paper award from ICCAS-SICE 2009. He can be contacted at email: nhutdt@uit.edu.vn.

Hoai-Nhan Nguyen received the B.S. degree from the Department of Automatic Control Engineering and Manufacturing Automation, Faculty of Mechanical Engineering of Ho Chi Minh City University of Technology, Ho Chi Minh city, Viet Nam, in 2004. He received the Ph.D. degree in School of Electrical Engineering, University of Ulsan, Ulsan city, South Korea in 2014. His current research interests are Robotic manipulator calibration, collaborative robots, Embedded control system for robots, mobile robotic assistant. He can be contacted at email: nhannh@uit.edu.vn.

Minh-Son Nguyen received B.Engr. and M.Engr. in Computer Engineering from HCMCity University of Technology, VietNam in 2001 and 2005 respectively and received Ph.D degree in Electrical Engineering at The University of Ulsan, Korea in 2010. In currently, Dr. NGUYEN is Dean of Faculty of Computer Engineering at University of Information Technology-VietNam National University at HCMCity. Dr. NGUYEN is also Director of Automotive R&D LAB of UIT and Member of Committee of Science and Technology of SaiGon High-Tech Park. His research is focussed on fields of wireless embedded internet, AI for IoT, smart system and System-on-Chip Design. He can be contacted at email: sonnm@uit.edu.vn.