2020 IEEE 16th International Conference on Control & Automation (ICCA)
Oct 9-11, 2020 (Virtual) Sapporo, Hokkaido, Japan

# A Novel Intelligent Robot for Epidemic Identification and Prevention

Hongrui Sang, Zhipeng Wang, Bin He, Member, IEEE, Yanming Zhou, and Gang Li

*Abstract*— **In order to guard against the 2019-nCoV new coronavirus, this paper proposes a novel intelligent robot for epidemic identification and prevention. To examine people's health more comprehensively, an infrared thermal imaging camera is equipped on the humanoid eyes-head robot platform that has the ability to hear, speak and watch. The robot not only has common capabilities such as face recognition, cough detection, mask detection, and heart rate detection, but also epidemic prevention and detection capabilities such as respiratory frequency and body temperature detection.**

## I. Introduction

The new coronavirus (2019-nCoV) is a new virus that can cause respiratory diseases in humans and can be transmitted from person to person. Its outbreak has a major impact on global health [1]. Designing and constructing an epidemic prevention and control robot with autonomous detection capabilities is a very significant task. This can not only improve the management efficiency of public places, but also reduce personnel contact and avoid cross infection [2]. This requires the robot to have high detection efficiency and good autonomy. Therefore, the perception system of such robots not only needs to have the ability to quickly recognize various respiratory symptoms, but also needs to provide human-computer interaction capabilities to autonomously complete multi-person detection tasks [3]. Our goal is to design and construct a detection system that achieves these capabilities, and study the higher-level development of the system in terms of autonomous epidemic prevention and control. In the next section, we will specify components of the robot. Section III details the ability of system identification and detection for epidemic prevention and control. Practical application performance of the robot is presented in section IV.

## II. SPECIFICATION OF THE ROBOT

We propose a new type of robot with epidemic prevention and detection capabilities (see Figure 1). With the advantages of the eyes-head [4][5] humanoid head, the robot can make good use of interactive services and cognitive vision research. We try to use works and turn them into a practical application. In the next, we will provide the developed eyes-head platform in detail about kinematics and motor, sensor, computation system.

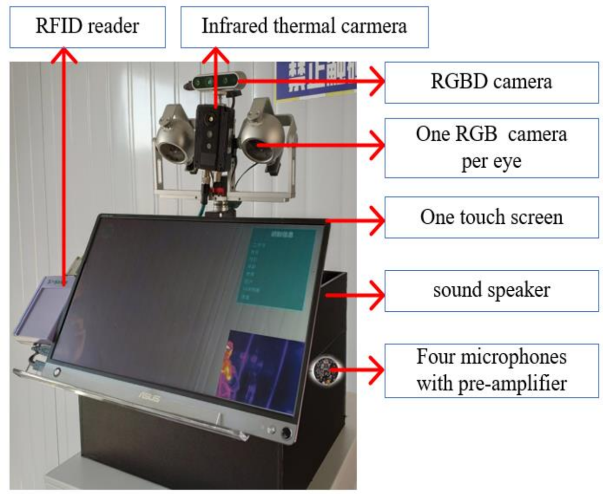


Figure.1 The intelligent robot for epidemic identification and prevention.

### A. Kinematics

Human-like neck-eye system has a very complex kinematic structure [6]. Simple mechanical connections are not enough to represent physiological details [7]. However, we will not copy these details completely, but rather to represent the anthropomorphic kinematics through mechanical movement. The head in our design contains seven DOF ($\theta_1$, …, $\theta_7$) that can capture the essence of human head movements, as shown in Figure 2. The neck has three DOF ($\theta_1$, $\theta_2$, $\theta_3$). The head vision system has four DOF ($\theta_4$, …, $\theta_7$), and each eye has two DOF, eyes yaw ($\theta_5$， $\theta_6$), and eyes pitch ($\theta_4$, $\theta_7$). Each DOF of link is driven by a bus servo.

The servo driver is connected to the signal conversion board by bus. As shown in Figure 2, the conversion board can convert the serial bus of SCS15, SCS009 and RS485 bus of SM85CL to USB, which is convenient for the control unit to output control instructions centrally.

### B. Motor system

We select different motors for various structural parts, due to the different load capacity of the servo. All motors are bus-type, which can not only simplify the design of the motor driver, but also make full use of the characteristics of the bus, to make the motor system clearer and less wired. For the pitch($\theta_1$) and roll($\theta_2$) of nick, we use FEETECH Serial Control RS485 Servo SM85CL, which has a weight of 140g, a

The work was supported by National Key R&D Program of China (Grant No. 2018YFB1305304), National Natural Science Foundation of China (Grant Nos. U1713215, 61825303 and 51975415), Major Project of Special Development Fund for Shanghai Zhangjiang National Independent Innovation Demonstration Zone (Grant No. ZJ2019-ZD-003), Shanghai Science and Technology Commission Project (Grant Nos. 18DZ1205706, 18DZ1202703). (*Corresponding author: Bin he*)

Hongrui Sang, Zhipeng Wang, Bin He, Yanming Zhou, and Gang Li are all with the Department of Control Science & Engineering, Tongji University, Shanghai 201804, China(e-mail: jasonsang@tongji.edu.cn; wangzhipeng@tongji.edu.cn;hebin@tongji.edu.cn;yanmin.zhou@tongji.edu.cn, lig@tongji.edu.cn )

Table.1 The intelligent robot for epidemic identification and prevention.

| Table Head | Servo specification | | | | | |
|---|---|---|---|---|---|---|
| | ***Size*** | ***Weight*** | ***Stalled torque*** | ***No-load speed*** | ***Driving voltage*** | ***Max angle*** |
| SM85CL | 62mm*34mm*47mm | 140g | 85kg·cm | 0.27sec/60° | 12V | unlimited |
| SCS15 | 40mm*20mm*40.5mm | 56g | 15kg·cm | 0.18sec/60° | 6V | 300° |
| SCS009 | 23.2mm*12mm*25.5mm | 12.5g | 2.3kg·cm | 0.27sec/60° | 6V | 280° |

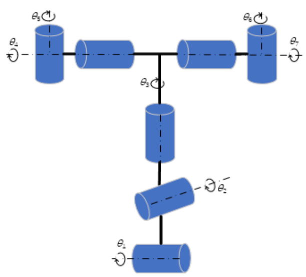

Figure.2 The kinematics of the head with seven DoF.

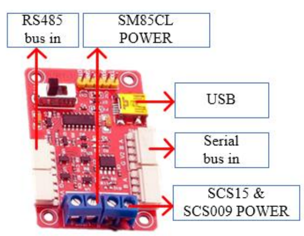


Figure.3 Signal conversion board for servo bus.

stall torque of 85kg·cm and a position resolution of 0.088 °. For the yaw ($\theta_3$) of neck, we use FEETECH Serial Control BUS Servo SCS15, which has a weight of 56g, a stall torque of 15kg·cm, and a position resolution of 0.29 °. For the steering system of the vision system, FEETECH Serial Control Servo SCS009 is selected, which has a weight of 12.5g, a stall torque of 1.89kg·cm, and a position resolution of 0.27 °. Refer to Table 1 for detailed parameters of the servo.

The servo driver is connected to the signal conversion board by bus. As shown in Figure 3, the conversion board can convert the serial bus of SCS15, SCS009 and RS485 bus of SM85CL to USB, which is convenient for the control unit to output control instructions centrally.

### *C. Perception system*

**Vision system:** It is necessary to first recognize the presence of face objects to perform the task of multi-person autonomous detection, and then analyze the necessity and order of detection of these objects. Although the first task is closely related to visual search and can be done with RGB cameras [8], it is difficult to analyze human behavior from the video stream [9], so we choose the Intel RealSense D435 depth camera, through which we can get the depth information of the faces in the image videos or pictures that is used as part of our primary and secondary sequence judgment. At the same time, live detection can be performed to prevent photo fraud [10]. For the intraocular camera, we use a wide dynamic camera with no distortion lens.

For temperature related detection tasks, we need to complete average face temperature measurement and breath detection based on heat map changes. We use the IRAY A8 infrared thermal imaging camera, which is a card-type multi-point non-contact infrared temperature measuring camera. It has a frame rate of 25fps, includes an infrared detector with a resolution of 240 * 180, and a field of view of 33.4 ° * 25.4 °, temperature measurement range is -20 ° C ~ 120 ° C, supports emissivity correction from 0.01 to 1.0 and temperature measurement accuracy ± 0.2 ° C.

**Sound System:** The platform is equipped with a four-channel microphone system for detection and localization of sound events. We choose the Respeaker Mic array v2.0 far-field microphone array which directly supports USB Audio Class 1.0 (UAC 1.0). All the major operating systems, such as Windows, MacOS, and Linux, are compatible with UAC 1.0, so Respeaker can run as a sound card and can easily implement audio processing algorithms such as Voice Activity Detection (VAD), Direction of Arrival (DOA), Beamforming, and Noise Suppression (NS). In addition, a speaker is connected to the Respeaker as a voice output device.

**RFID System:** As many of the existing access management systems are controlled by RFID, in order to be compatible with them, our robot is also equipped with an RFID reader, which is controlled by a separate computing unit to ensure the security of the internal information system.

### D. Computing System

The computing system contains three main computing units, NUC8 I7HNK, Jetson AGX Xavier, and Raspberry Pi 4b. Through this combination, our system has great flexibility. The NUC8 I7HNK is a mini PC from INTEL, it has an independent graphics card and a 1GHz main frequency in a volume of 1.2L. We have it equipped with Debian Linux to realize the eye-head controller based on ROS, display of human-computer interaction information and maintenance of message queue for detection system. In addition to the CPU and GPU, the Jetson AGX Xavier is also designed with the new Deep Learning Accelerator (DLA) and Programmable Vision Accelerator (PVA), among which the DLA is a new dedicated unit for machine reasoning acceleration with the INT8 computing performance up to 11.4Tops. After equipped with CentOS, it is used to perform face capture tasks locally and perform face search, face attribute detection, mask wearing detection and other tasks remotely by accessing

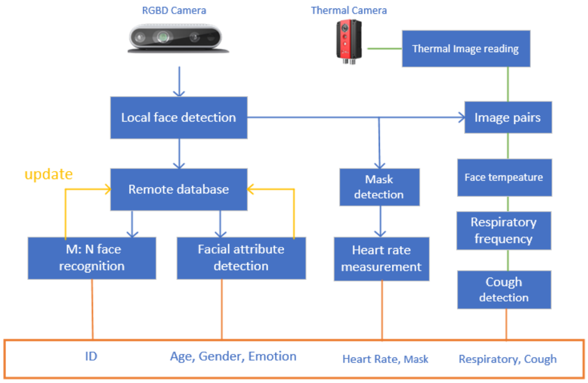


Figure.4 The overview of vision system task.

Figure.5 Thermal image during inhalation (left) and exhalation (right).

remote computing. The Raspberry PI 4b is a tiny computer in the size of a credit card and we use it to handle tasks related to RFID. All devices are connected via Ethernet and communicate using a network.

## III. VISION TASKS

An overview of the software for computer vision task is shown in Figure 4. The computer vision task is divided into three parts: local image acquisition and processing, remote image search and attribute detection, and heat map processing.

### *A. RGBD image acquisition and processing*

RGBD image acquisition and processing: The system acquires depth images from the RGBD camera and completes the face capture task locally. It obtains the pixel coordinates and depth information of the face in each frame. Moreover, it detects whether a person wear a mask by object detection and completes heart rate measurement by detecting the changes in facial skin color caused by blood circulation.

### *B. Remote image retrieval and facial attribute detection*

Remote image calculation contains face database, M: N face recognition and facial attribute detection including age, gender, and expression. Furthermore, to improve the accuracy of detection, the system updates the latest face database by the results of image retrieval and facial attribute detection.

### *C. Thermal image acquisition and processing*

The system acquires thermal images from the thermal camera and pairs with depth images to obtain the corresponding thermal image position of each face position. It detects the average temperature of the face by calculating the weighted average temperature of each pixel. In addition, it can detect the respiratory frequency and cough by detecting the changes in facial mask heat map with time, as shown in Figure 5.

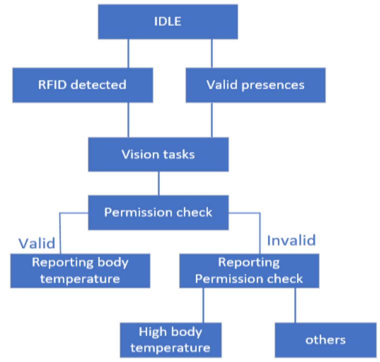


Figure.6 The overview of vision system task.

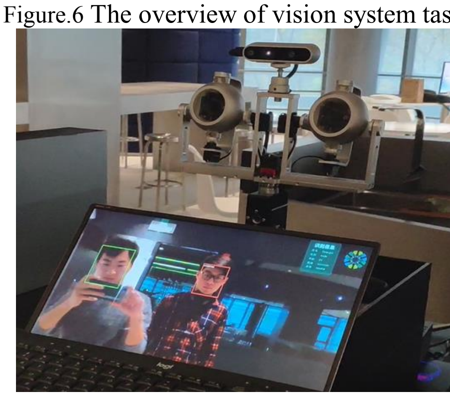

Figure.7 The red rectangle is the selected tracking face.


Figure.8 Working scenario of the system.

## IV. RESULTS

The task flow diagram of the system is shown in Figure 6. When it is in waiting to work, the head eyes are naturally retracted downward. When someone swipes a card or appears within the effective temperature measurement distance (less than 1m), it will activate other detections. After completing the face detection and identification, the head eyes will raise and track according to the position of the face, the tracking face depends on the face recognition score and exclude the examined people , as shown in Figure 7, the tracking speed of the servo is set to 100 ° / second. At the same time, the display interface will update the person's identification information, and whenever other facial attribute detection information is updated, the display interface will also be updated synchronously. When the detection is completed, the person’s body temperature and whether to be allowed to pass will be announced by voice. When the temperature and personal information meet the detection requirements, the person’s temperature and being allowed to pass will be announced. When the passage conditions are not met, a warning of high temperature or other reasons will be announced. The entire detection process including heart rate, mask detection, temperature measurement, respiratory frequency, and cough can be completed within 15 second.

In practical application, our robot is placed in the temporary access detection channel at the entrance of the school. This is a semi-enclosed environment. The working scenario of system is shown in Figure 8.

## V. CONCLUSION

This work proposes a novel epidemic prevention detection robot based on eye-head human-like head platform. We give an introduction of the detection system, vision method and detection work flow in detail. The results show that the system is able to successfully identify and track faces, detect cough, mask wearing, heart rate, respiratory frequency and body temperature, etc. And it can interact with people through tracking, voice and display ability so as to independently complete the detection task.

As future work, we will explore the detection ability of the system under different conditions of light, temperature, humidity, and gradually improve the adaptive ability of the system to the environment.

## ACKNOWLEDGMENT

The work was supported by National Key R&D Program of China (Grant No. 2018YFB1305304), National Natural Science Foundation of China (Grant Nos. U1713215, 61825303 and 51975415), Major Project of Special Development Fund for Shanghai Zhangjiang National Independent Innovation Demonstration Zone (Grant No. ZJ2019-ZD-003), Shanghai Science and Technology Commission Project (Grant Nos. 18DZ1205706, 18DZ1202703).

## REFERENCES

[1] World Health Organization. "Novel Coronavirus ( 2019-nCoV): situation report, 3." (2020).

[2] Wu, Joseph T., Kathy Leung, and Gabriel M. Leung. "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study." The Lancet (2020).

[3] Katzenmaier, Michael, Rainer Stiefelhagen, and Tanja Schultz. "Identifying the addressee in human-human-robot interactions based on head pose and speech." Proceedings of the 6th international conference on Multimodal interfaces. 2004.

[4] Das, Sumit Kumar. "Realistic interaction with social robots via facial expressions and neck-eye coordination." (2015).

[5] Roncone, Alessandro, et al. "A Cartesian 6-DoF Gaze Controller for Humanoid Robots." Robotics: science and systems. Vol. 2016. 2016.

[6] Zatsiorsky, Vladimir M., and Vladimir M. Zaciorskij. Kinetics of human motion. Human Kinetics, 2002.

[7] Asfour, Tamim, et al. "The karlsruhe humanoid head." Humanoids 2008-8th IEEE-RAS International Conference on Humanoid Robots. IEEE, 2008.

[8] Welke, Kai, Tamim Asfour, and Rudiger Dillmann. "Active multi-view object search on a humanoid head." 2009 IEEE International Conference on Robotics and Automation. IEEE, 2009.

[9] Wu, Di, Nabin Sharma, and Michael Blumenstein. "Recent advances in video-based human action recognition using deep learning: a review." 2017 International Joint Conference on Neural Networks (IJCNN). IEEE, 2017.

[10] Atoum, Yousef, et al. "Face anti-spoofing using patch and depth-based CNNs." 2017 IEEE International Joint Conference on Biometrics (IJCB). IEEE, 2017.