IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 4, 2021, accepted May 17, 2021, date of publication May 20, 2021, date of current version June 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3082143

In-Robot Network Architectures for Humanoid Robots With Human Sensor and Motor Functions

CHENGYU CUI AND SUNGKWON PARK, (Senior Member, IEEE)
Department of Electronics and Computer Engineering, Hanyang University, Seoul 04763, South Korea

Corresponding author: Chengyu Cui (ccy041124@hanyang.ac.kr)

This work was supported by the National Research Foundation of Korea (NRF) of the Korean Government, Ministry of Education (MOE) under Grant 2018R1D1A1A02048970.

ABSTRACT The design concepts of the in-robot network (IRN) architectures of humanoid robots are proposed in this paper. First, this paper reveals the network requirements for humanoid robots to realize perception abilities and action execution abilities near those of human beings. Humanoid robots need to be equipped with many sensors to collect surrounding environmental information. It is also necessary to use many motion actuators to enhance the degree of freedom to achieve the smooth motion abilities of humans. To maintain reliable data transmission between a number of nodes, an efficient and reliable IRN architecture is needed. This paper first discusses the limitations of existing humanoid robots in the number of sensors and degrees of freedom and points out that one of the reasons for this limitation is the lack of reliable network architectures. In-vehicle networks (IVNs) and various network technologies are used. These include time-sensitive networks and control area networks (CANs) to name a few. Additionally, heterogeneous network protocols are used. IRNs also include networks of sensors and actuators with different performance parameters such as bandwidth, delay, and transmission speed. IRNs may adopt design ideas similar to those of IVNs to satisfy the network requirements for humanoids. To accomplish this, three feasible IRN architectures are proposed and analyzed. Finally, to compare and analyze the three proposed IRN architectures, this paper uses OMNeT^{++} simulation software.

INDEX TERMS Humanoid, in-robot network, sensor network, actuator network, backbone network, bandwidth, delay, bit rates.

I. INTRODUCTION

Humanoids are intelligent robots with appearances and functions similar to those of humans. This kind of robots pursues the high fidelity of human beings. Not only the facial and limb features but also the movement abilities and shapes of humanoids can highly emulate those of real human beings. Such humanoid robots can also act like human beings, have the same expression, and make corresponding reactions when encountering issues. The reason why humanoid robots can act like humans and have human behavior is that they have a central system constructed by sensors, which controls and directs the behavior of the robot in a way similar to that of the brain. To achieve such a function, various parts of humanoids are equipped with a variety of sensors, including optical sensors, which enable them to see surrounding objects and distinguish their sizes and colors; sound sensors, which enable the robots to hear surrounding sounds; and touch sensors, which give humanoids a tactile sense similar to that in human beings [1], [2]. Some humanoids are also equipped with ultrasonic sensors, so that they can hear ultrasonic waves that cannot be heard by human ears [3], [4]. However, in order to realize these complex skills, corresponding powerful network support is inevitably needed to carry the extensive amount of communication generated between thousands or even tens of thousands of sensors all over the body of the humanoids.

In this paper, we propose the in-robot network (IRN), which is intended to support the extensive amount of data communication generated between the sensors, processors, and actuators in humanoids. We will focus our research on the communication network issues rather than social functionalities such as voice communication and vision abilities. These issues include the network architecture to ensure reliable data transmission between communication elements with different network requirements for different functionalities. Although social functions are one of the indispensable capabilities of

The associate editor coordinating the review of this manuscript and approving it for publication was Junaid Shuja.

humanoid robots, the social-level information processing in the application layer is not included in the scope of this paper. However, we incorporated its data when studying the required amount of data and transmission speed, such as the bit rates of the audio streams and the data packet sizes of the actuator control commands sent by the processors when performing body movements. Similar to the neural network of the human body, the sensors in the humanoids play the role of sensory neurons and collect the temperature, humidity, pressure and other environmental factors of the external environment. These environmental factors are transmitted through the IRN to the central procession unit (CPU), which serves as the robot's brain. The processor analyzes these data and generates further commands to control the humanoids' actions. The actuators play the role of motor neurons, similar to the muscle tissue of the human body, enabling the humanoids to perform various complex motion functions similarly to humans.

In previous studies, we analyzed the complex perception of the human body from various perspectives such as temperature, pressure, vision, and hearing [5]. By investigating literature in the areas of human biology, we discussed the number of various sensory spots, sensitivities, and abilities to identify environmental factors of the human body. Based on the discussion, we suggest the distributions of the number of sensors and the sensitivities of the sensors in the sensor network for the humanoids. We first estimated the required data size for various sensors in the networks and the minimum communication delay requirements. We used these to calculate the network bandwidth requirements that need to be realized in the humanoid IRN. In a similar way, we then estimated the data size of the packets received by the actuators that act as human motor organs, data transmission speed, and communication delay requirements.

In this paper, we refer the concept of in-vehicle network (IVN) and apply it into IRN. Our goal is to design a set of IRN architecture forward-looking and apply it to the humanoid industry in the future. When the robot industry level in the future can restore human perception to a high degree, the IRN architecture we proposed can be directly applied to it to carry the huge network communication traffic and meet the basic communication requirements. We not only refer to the bandwidth and delay requirements estimated in previous studies but also propose three near-optimum humanoid IRN architectures through the viewpoints of the weight, cost, and wiring length of the IRN. These architectures under different requirements are evaluated through comparative analysis. To reduce the traffic loads of the backbone networks in these architectures and prevent data congestion, we divide the architectures into several different domains. The partitioning of the domains can be changed according to the different requirements of the robot function. We found that the existing research on IRNs is only at the basic ethernet and control area network (CAN) level and does not solve the large-scale communication network issues we proposed. [1], [6], [7] Therefore, we examined the IVN and found that in order to provide the required security level, time sensitivity, and communication speed for different network nodes, the IVN divides the entire vehicle network into several domains and assigns different network communication protocols. Therefore, we try to incorporate the concept of the IVN into and the IRN and propose a new architecture for IRNs that is different from that of IVNs.

Then, we used the OMNeT^{++} tool to simulate the three IRN architectures that we proposed through two kinds of scenarios to compare the performances of the three IRN architectures in fully loaded communication environments. The end-to-end delay was used as a measure of performance to compare the different network architectures. We tested the average propagation delay from sensors to the CPU and from the CPU to actuators.

The organization of this paper is as follows. Section II identifies and explains problems in IRNs of existing humanoids and proposes general requirements of humanoids from a biological perspective such as the numbers of sensors and actuators and required bandwidths. Section III introduces, explains and discusses the general ideas of the three IRN architectures we proposed. In this section, the possible weights, costs, and wire lengths of IRNs with the proposed architectures are also discussed. Section IV sorts out and compares the end-to-end delays for the different architectures. Finally, Section V concludes this paper.

II. BACKGROUND AND RELATED WORK

The IRNs of humanoids are generally divided into sensor networks, actuator networks, and the backbone network that carries all communication categories. Investigations show that existing humanoids are not yet able to realize all the functions of human beings [8]–[11]. Instead, a limited number of sensors and actuators are employed to realize part of the functions. Therefore, it is necessary to design an efficient IRN architecture to support the very large amount of traffic between increasing numbers of sensors and actuators. At the same time, it is also necessary to find the proper bandwidths, traffic engineering, and architectures to guarantee the minimum transmission delay requirements. In this section, we first introduce the functions, degrees of freedoms, and sensor distribution of existing humanoid robots and show that they do not have enough sensors and actuators. Therefore, in the second part of this section, the number of elements and minimum bandwidth requirements of sensor networks and actuator networks are introduced, which may serve as an important basis for the designs of the IRN architectures in this paper.

Humanoid Nadine, jointly developed in 2015 by Kokoro Japan and Nanyang Technological University, is a social robot that can communicate with people. Nadine has a human appearance with natural skin and hair. Nadine can have a continuous dialogue with humans by recognizing facial expressions and dialogue content. Nonetheless, a temperature sensor that senses the ambient temperature and a pressure sensor that achieves tactile ability are not included in order to achieve these social functions. Therefore, Nadine still lacks

the ability of human skin to perceive the outside world. Nadine has 27 degrees of freedom, which supports only facial expressions and upper limb movements [8].

Sophia is a social humanoid robot developed in 2016 by the Hong Kong-based company Hanson Robotics. The camera in Sophia's eyes is combined with computer algorithms to enable her to see. She can follow faces, maintain eye contact, and identify individuals. In addition, Sophia has the ability to imitate more than 60 facial expressions [9]. However, as Sophia is a social robot similar to Nadine, the lack of action actuators makes it impossible to fully restore human motion functions.

The humanoid ASIMO developed in 2011 by Honda can talk to multiple people at the same time. When it encounters other people in action, it will predict the direction and speed of the opponent and precalculate alternative routes to avoid collisions. The exercise capacity and range of motion of the legs enable not only walking, running, and walking backwards but also jumping on one foot and two feet, changing directions while jumping, and walking on slightly uneven ground [10]. ASIMO 2011 has a walking speed of 0.75 m/s, a running speed of 2.5 m/s, and a total of 57 degrees of freedom.

Atlas, developed in 2013 by the American-based robotics company Boston Dynamics, is specially designed for mobile operation. It is very good at walking on a wide range of terrains, including snow, and can perform backflip and rollover maneuvers. The whole body contains 28 joints that are actuated hydraulically [11]. In addition, it can adjust the dynamic interaction between the whole body and the external environment through algorithms for motion planning.

In the in-vehicle network (IVN), the entire network structure is divided into four or five domains according to the differences in traffic information and functions, including the infotainment, body, power train (PT)/chassis, and advanced driving assistant system (ADAS) domains [12]. The IVN connects the various electronic control units by means of different communication media, so that the information of the sensors and actuators of each electronic control system is shared among the various units through multiplex transmission technology. In the IRN, different body parts have sensors with different accuracy and sensitivity levels. Therefore, it is necessary to allocate different communication protocols in a way similar to that in the IVN in order to reduce communication costs while providing ideal data transmission capabilities. Similarly, referring to the IVN, we divide IRN architectures into several different domains according to different functional requirements and physical location factors. This can not only divide network devices to meet similar bandwidth and delay requirements but also effectively reduce the complexity of network wiring and at the same time reduce the weight and economic costs of the entire structure. In addition, in the IVN, according to the distinction between functions and safety requirements, respective priority principles are formulated for different nodes in each network domain to obtain the timely response required in a vehicle collision accident. In the IRN, network components also need to be divided in a similar way. For example, the actuators on the torso and upper limbs of a humanoid robot are distinguished from those on the lower limbs in terms of function and use frequency. Moreover, in situations where humanoid robots may lose their balance, lower limbs often require more precise movement control than upper limbs. Therefore, in certain scenarios, the data transmission process generated to control the movement of the lower limbs often generates more traffic and requires a higher priority and less transmission delay. Therefore, for different domains, it is necessary to match the adapted network protocol to the requirements for data transmission rate, bandwidth, network priority judgment, and time sensitivity abilities.

The specific network protocol division can refer to the in-vehicle network. In the IVN, different types of nodes are respectively connected by multiple buses of different speeds, and gateway services are used to realize the information sharing and network management of the entire vehicle. Among them, equipment such as windows, seats, and lighting systems use the low-speed network connection of the local interconnect network (LIN) protocol [13]. This type of equipment has low requirements for network transmission rates, generally less than 10 kbps. For fault diagnosis, air conditioners and instrument panels require a higher transmission rate but do not require high real-time performance, so most of them use low-speed CAN protocol networks for connection [14]. For engine control, anti-lock brake systems (ABSs), electronic stability programs (ESPs), and suspension control, a high-speed CAN protocol and FlexRay protocol with a transmission rate of 125∼1000 kbps are used between devices that have high real-time requirements and directly affect driving safety. Technologies such as media oriented systems transport (MOST), Ethernet, and Bluetooth are used in multimedia and navigation systems that require high transmission rates [15]–[17]. In addition, the IEEE P802.3cg Single Pair Ethernet Task Force proposed by the 802.3 WG-Ethernet Working Group in 2019 can support data transmission and power supply at the same time with a single pair of cables [18]. The common Ethernet before these requires two pairs of cables and separate cables for the power supply. It can also greatly reduce costs and reduce the weight of the equipment while providing 10 Mbps of data signaling. The transmission rates of 100, 200, and 400 Gbps provided by IEEE P802.3ck can be applied to the network used by the camera equipment that generates a large amount of data transmission in humanoid robots [19].

In addition, when building the IRN of humanoid robots, not only are the bandwidth and delay of the network important indicators for measuring network performance but also the actual physical weight and cost of the entire network equipment are factors that cannot be ignored. We transformed the above performance factors into equation 1 to compare the performances of networks of different architectures. Equation 1 shows the cost function to be minimized with the major parameters to be adjusted in designing an IRN architecture.

The parameters in equation 1 are dependent on the market but show that they are the key parameters to be considered.

$$P = \alpha_1 \left[\sum_{i=0}^{N_s-1} W_S(i) + \sum_{j=0}^{N_n-1} W_N(j) + LW_L \right] + \alpha_2 \left[\sum_{i=0}^{N_s-1} C_S(i) + \sum_{j=0}^{N_n-1} C_N(j) LC_L \right] \alpha_3 \frac{D_a}{B_S}, \quad (1)$$

where W_S (i) and W_N (j) represent the actual physical weights of the i^{th} switch and j^{th} node, respectively. W_L represents the physical weight of the unit length. L represents the total length of the network link. C_S (i) and C_N (j) represent the costs of the i^{th} switch and j^{th} node, respectively. C_L represents the cost of the network link. B_S is the required bandwidth of the backbone network, and D_a is the average propagation delay during packet transfer process in the IRN. The values of weights α_1, α_2, and α_3 can be adjusted according to the relative emphasis on the weight, cost, bandwidth, and delay of the network elements. Additionally, $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

III. IRN ARCHITECTURE CONCEPT

Based on the numbers of different sensors and actuators and bit rates for an ideal humanoid generated during the communication process provided, we analyze network load situations and divide the networks into several domains as in the concept of humanoid domains. This section introduces three IRN architectures designed to meet the requirements for different communication scenarios.

A. IRN ARCHITECTURE DIVIDED BY FUNCTIONALITY

To design various IRN architectures for a humanoid, we first divide all sensor networks and actuator networks into different domains based on functionalities and physical locations. The purpose of designing the IRN of the robot with domains is to greatly reduce the load of the backbone network to prevent possible data conflicts during the communication process. The basic architecture of the humanoid we proposed is designed based on the star topology. In addition, different domains require different network bandwidths for economic reasons, and different domains can be equipped with different transmission methods. For example, some lower-priority networks can be equipped with LIN or CAN protocols. Some domains with larger traffic can use 802.3cg for a 10 Mbps link or 802.3ch for 2.5, 5, and 10 Gbps links [18], [20].

The sensor network contains nearly 700,000 sensors, and these sensors need to transmit a very large amount of uplink data to the central processing unit every second. After the central processor analyzes and calculates the received data, it transmits the downlink commands to the actuators that are all over the humanoid body. The cameras have a bit rate of up to 8.294 Gbps, which occupies 83.02% of all communication. Therefore, we group it into the head domain together with the smell sensor and the mic. This is because the smell sensor generates a bit rate of only 4.15 kbps, which is the least among all sensors. In addition, the smell sensor, mic, and camera are usually placed on the head of the robot. Grouping them into the head domain can also reduce the wiring length and weight of the robot. Correspondingly, the actuators in the face and mouth are grouped into the head domain and share a switch with the smell sensors, mics, speaker, and cameras. They are then connected to the backbone network through the switch.

In the head domain, the bandwidth required by the communication path connected to the backbone network can be calculated by equation 2,

$$B_S^H = 2B_c + 2B_m + B_o + N_a^H \cdot B_a, \quad (2)$$

where superscript H represents the head domain, and B_S^H is the required bandwidth of the head domain. B_c is the bandwidth required by a camera. B_m is the one required by each of the two mics. B_o is the total bandwidth required by the smell sensor. N_a^H is the total number of actuators in the head domain, and B_a is the bandwidth each actuator requires.

The temperature sensors and pressure sensors in the other domains total more than 700,000. According to difference in physical locations, we divide these sensors all over the humanoid robots into three subdomains and connect them to a common switch by a star topology. These three subdomains coexist in the skin domain and share a switch to connect to the backbone network. We divide them into the left arm and hand subdomain (LAH subdomain), right arm and hand subdomain (RAH subdomain), and leg, foot, torso and head subdomain (LFTH subdomain) based on the number of sensors in different parts. Among them, the LAH subdomain and the RAH subdomain include temperature and pressure sensors, respectively, on the robot hand, forearm, and big arm, totaling 813,714. The LFTH subdomain contains 519,603 temperature and pressure sensors located on the leg, foot, torso, and head of the robot.

Therefore, the total bandwidth required by the communication link responsible for transmitting all sensor data to the backbone network in the skin domain can be derived from equation 3,

$$B_S^S = N_t \cdot B_t + N_p \cdot B_p, \quad (3)$$

where B_S^S is the required bandwidth of the skin domain, which can be obtained by the sum of the product of total number of temperature sensors N_t in the skin domain, the required bandwidth of temperature sensors B_t, the product of the total number of pressure sensors N_p and the required bandwidth of pressure sensors B_p.

For the actuator part, we consider that there are differences in link length due to the position of the actuator, and in some cases, part of the IRN may be physically damaged due to external forces such as unintentional impact. To deal with this situation, each network line has a corresponding back-up path. At the same time, even in the case of failure, the humanoid robots still need to maintain the most basic movement ability. Therefore, we divide the leg and torso parts responsible for moving and the arm part into different domains to block the impact from other parts under failure situations. So, the actuator network is divided into the arm and

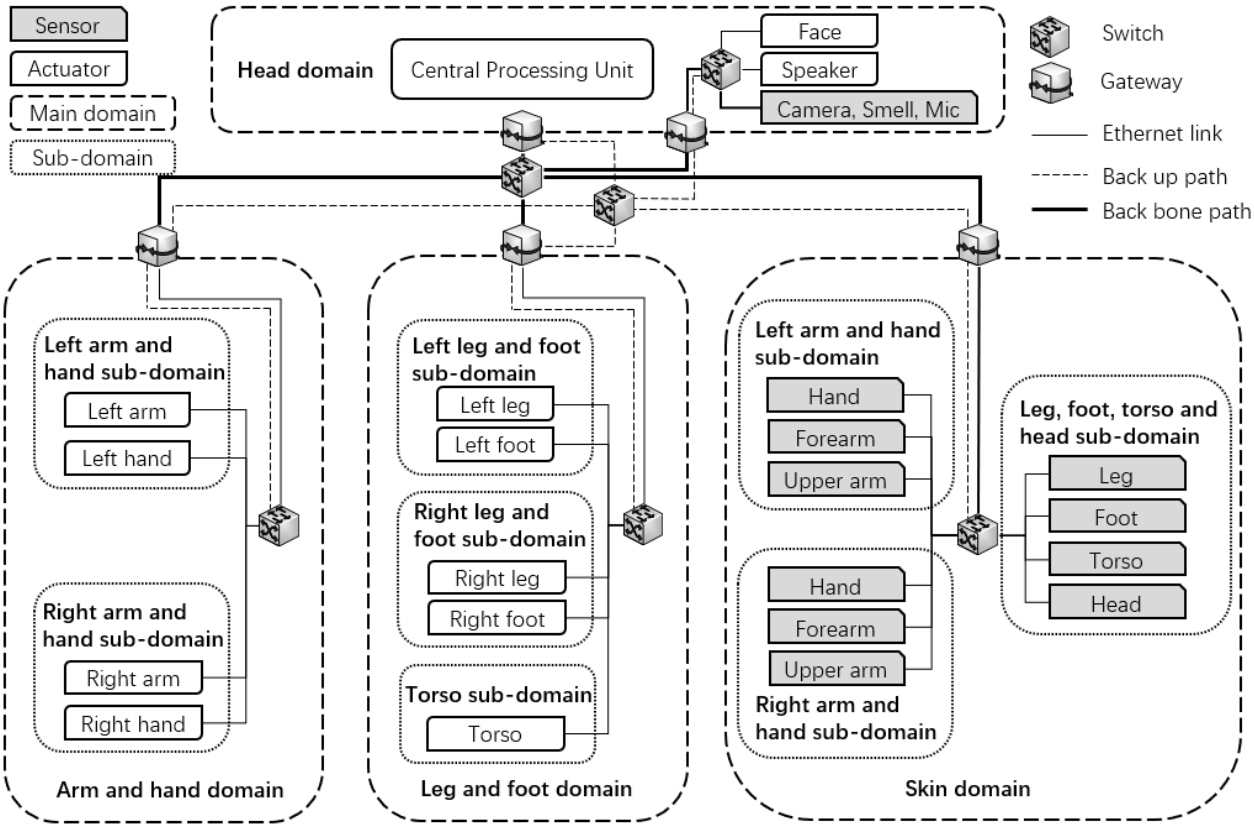

FIGURE 1. Architecture of the IRN domain divided by functionality. The architecture includes the head domain, arm and hand domain, leg and foot domain, and skin domain. All domains connect to the backbone network and share one switch. It also includes a backup switch on the backbone for use in the event of a single node failure.

hand domain on the upper body and the leg and foot domain on the lower body. The actuators of the torso part of the robot are grouped into the leg and foot domains. These two domains are further divided into left and right subdomains. The arm and hand domains contain 44 actuators and generate a bit rate of 18.48 Mbps. The bandwidth requirement B_S^{AH} can be calculated by equation 4,

$$B_S^{AH} = N_a^{AH} \cdot B_a, \tag{4}$$

where B_S^{AH} is the required bandwidth of the arm and hand domains. In addition, N_a^{AH} is the total number of the actuators in the arm and hand domains. B_a is the required bandwidth of an actuator.

The leg and foot domains contain 104 actuators and generate a bit rate of 43.68 Mbps. The required bandwidth of the leg and foot domains can be calculated by equation 5,

$$B_S^{LF} = N_a^{LF} \cdot B_a, \tag{5}$$

where B_S^{LF} is the required bandwidth of the leg and foot domains. The N_a^{LF} is the total number of the actuators in the leg and foot domains. In addition, a switch is allocated to each of the two domains to connect to the backbone network. Therefore, the required total bandwidth B_S of the backbone network can be represented by equation 6,

$$B_S = B_S^H + B_S^{AH} + B_S^{LF} + B_S^S, \tag{6}$$

In Figure 1, the different domains are marked with dashed boxes. Since different domains produce different bit rates, different levels of network bandwidth need to be allocated. The link between the switch and the backbone in the arm and hand domain and leg and foot domain needs to use a bandwidth of 200 Mbps. The link between the switch and the backbone in the skin domain needs a bandwidth of 2 Gbps. In addition, because the head domain contains a high bit rate video transmission link, it needs to use 10 Gbps Ethernet. All domains are connected to the backbone through the corresponding gateway and share one switch. Considering that a single node failure may occur in the star topology, the impact on data transmission is fatal. Therefore, we believe that a backup switch should be added to the backbone to connect to all domains to provide backup paths for all domains in the event of a single node failure. Due to the differences in the characteristics of sensors and actuators in different domains, there will be differences in the priority of network use, time sensitivity, and requirements for fault tolerance mechanisms. Therefore, different domains should be assigned corresponding network communication protocols, so we set up a

TABLE 1. Bandwidth distribution in the skin domain of IRN Architecture divided by functionality.

Domain	Bandwidth Allocation
Head domain	8.295 Gbps
Skin domain	1.692 Gbps
Arm and hand domain	18.48 Mbps
Leg and hoot domain	43.68 Mbps

gateway on the link connecting each domain to the backbone. The bandwidth allocation required by the connection lines between each domain and the backbone network is shown in Table 1.

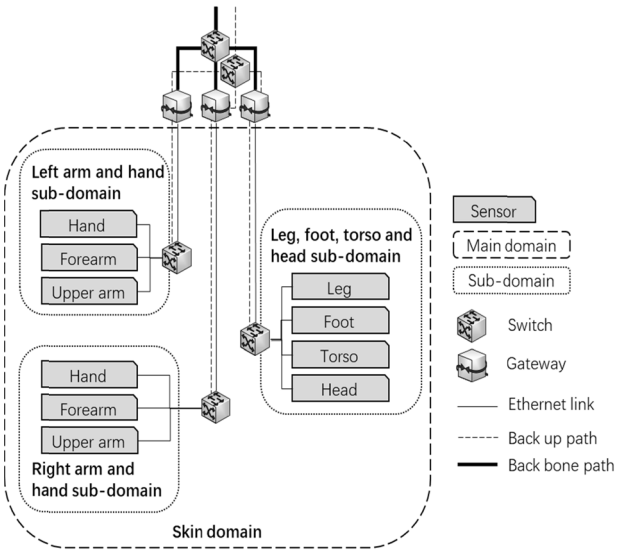

FIGURE 2. Skin domain of the switch load-minimized IRN Architecture. Each subdomain in the skin domain has a corresponding switch and gateway, which are individually connected to the backbone network, and has a separate backup path to connect to the secondary switch of the backbone.

B. IRN ARCHITECTURE WITH A MINIMIZED SWITCH LOAD

In the first IRN architecture of the humanoid, it can be found that the main communication traffic is concentrated in the head domain line that transmits camera images and the skin domain line that transmits temperature and pressure sensor data. The camera image is transmitted to the backbone network using a single line in the form of an uninterrupted data stream. However, because the number of sensors in the skin domain is too large and the demand for bandwidth is close to the 2 Gbps level, only assigning one switch to the skin domain may have a serious impact due to data transmission congestion. Therefore, in the second IRN architecture, each subdomain in the skin domain is equipped with a corresponding switch and connected to the backbone network through its own gateway, as shown in Figure 2. As a result, the network bandwidth requirements carried by the switch originally in the skin domain can be reduced by 2 Gbps. However, it is inevitable that due to the increase in switches, the complexity of the circuit and the weight of the robot will increase.

The network bandwidth allocated by the switches of the three subdomains in the skin domain is shown in Table 2.

TABLE 2. Bandwidth distribution in the skin domain of the switch load-minimized IRN Architecture.

Domain	Bandwidth Allocation
LAH subdomain	207.248 Mbps
RAH subdomain	207.248 Mbps
LFTH subdomain	1.277 Gbps

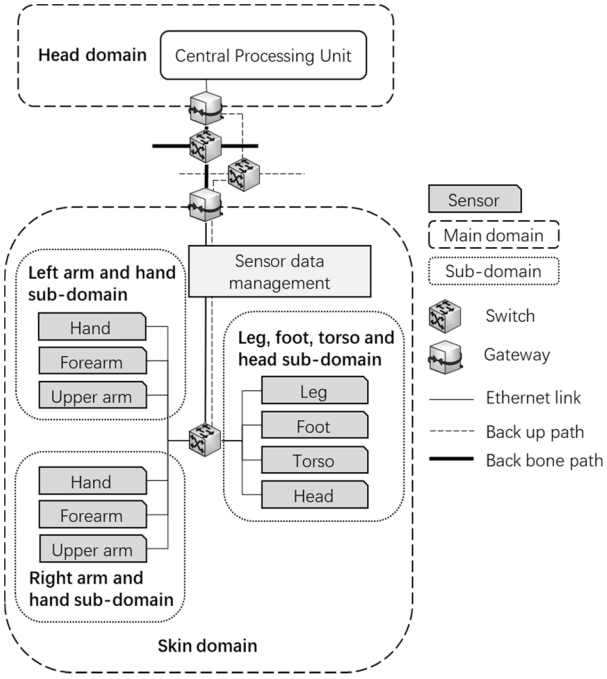

FIGURE 3. Skin domain of communication the traffic-minimized IRN Architecture. Sensor data management is added to the link between the skin domain switch and the backbone switch, which is responsible for filtering redundant data.

C. IRN ARCHITECTURE WITH COMMUNICATION TRAFFIC

While increasing the number of switches in high-traffic domains, the load of each switch can be reduced to a certain extent. However, this degree of improvement is still limited. Therefore, we believe that directly reducing the amount of communication is also an effective way to reduce the risk of communication congestion. We also carried out further transformations on the network based on the first architecture, as shown in Figure 3. In the third architecture, we added sensor data management to the line connecting the switch to the backbone network in the skin domain. Sensor data management is used to filter unnecessary data in the communication process to reduce the occurrence of redundant communication. For example, in a constant-temperature environment,

the temperature change sensed by the temperature sensor in the humanoid body is very small. At this time, if repeated values are always transmitted from each temperature sensor, it will inevitably cause a great load on the entire network. A threshold can be set in sensor data management, for example, if the temperature change value of the sensor output is less than 1 degree Celsius, the corresponding data can be filtered out, and a smaller data packet can be transmitted to the central processing unit to inform that there is no significant data variety. In this way, unless a large temperature change occurs during an emergency, under normal circumstances, the continuous transmission of repeated data can be reduced. When there is a sudden change in temperature and pressure, sensor data management then transfers the collected data to the central processing unit, so that the central processing unit can analyze the data and generate a timely response. Under ideal conditions, most of the time, the load of the skin domain switch can be reduced from the original 2 Gbps to below 10 Mbps.

To compare the three IRN architectures we proposed intuitively, we can use function 1 to measure the three architectures in terms of robot weight, cost and bandwidth. Since the switches and sensors that can be found on the market today cannot fully meet the network requirements proposed in this article, the calculation results are only used as a reference for comparing the three network architectures. We choose the TL-SG116 switch for the 200 Mbps link. The reference price of TL-SG116 D is $59.99, and its weight is 2.07 pounds. For the 10 Gbps link, we used the GS110EMX switch. Its price is $249.00, and its weight is 1.65 pounds. We chose SVPRO 5 MP as the camera for the robot's eyes. The price is $47.99, and the weight is 3.52 ounces. The microphone module uses the SUBALIGU 3PCS microphone sensor with a price of $11.98 and a weight of 0.81 ounces. We chose Lithonia Lighting 4000K as the speaker output device. The price is $75.00, and the weight is 1.95 pounds. In addition, we chose a HiLetgo 5pcs DHT11 temperature humidity sensor with a unit price of $10.49 and a weight of 0.63 ounces as the temperature sensor. For the pressure sensor, we chose the KOOBOOK 5Pcs GY-BMP280-3.3 pressure sensor module with a price of $6.59 and a weight of 0.63 ounces. For the actuators, we choose PA-07-1-5, with a unit price of $69.99 and a weight of 2.46 ounces. In addition, we chose optical fiber cables with a unit price of $2.33 per meter and weight of approximately 0.27 ounces/m use for network wire. The performance parameter results of the three IRN architectures calculated by function 1 are shown in Table 3.

Through the comparison, it can be found that compared with the first architecture, the second architecture is equipped with two more 200 Mbps Ethernet switches, and the third architecture is equipped with an additional CPU for the role of sensor data management. Therefore, the weight and cost efficiency are sacrificed, and the wiring complexity of the second architecture is higher than that of the other two architectures. The third architecture has the ability to filter redundant data packets due to sensor data management, so it can provide better delay performance in certain environments. Therefore, from the perspectives of bandwidth and delay performance, it can be found that the calculation results for the third architecture are better than those of the other two architectures. The calculation result for architecture one is more balanced than that of the other two.

TABLE 3. IRN Architecture performance results.

Architecture	Performance Calculation Result
Architecture 1	$30{,}367.21\alpha_1+5{,}701{,}030.68\alpha_2+301.47\alpha_3$
Architecture 2	$30{,}371.35\alpha_1+5{,}701{,}580.65\alpha_2+301.47\alpha_3$
Architecture 3	$30{,}367.4\alpha_1+5{,}701{,}890.66\alpha_2+(180.72\sim301.47)\alpha_3$

IV. SIMULATION RESULT ANALYSIS

Due to the huge scale of our simulation models, we choose OMNeT^{++}, which is relatively stable and can flexibly track and debug network parameters, as our simulation experiment tool. At the same time, we used and extended the network protocols provided in the open-source model library INET Framework in OMNeT^{++}. The simulation experiment of the IRN of the three humanoid robots is carried out using OMNeT^{++}. Taking the IRN architecture of a balanced bandwidth distribution robot as an example, the simulation props are first divided into four domains and connected in a star topology through the switch on the backbone network, as shown in Figure 4. Figure 5(a) shows the internal structure of the head domain, which contains the camera, mic, smell sensor and multiple actuators located in the head. These nodes are connected to the CPU through the switch located in the head domain. Figure 5(b) shows the skin domain of the IRN. The temperature and pressure sensors all over the robot are connected to the switch in the skin domain. Figure 6 shows the arm and hand domain and leg and foot domain containing the body part actuators of the IRN.

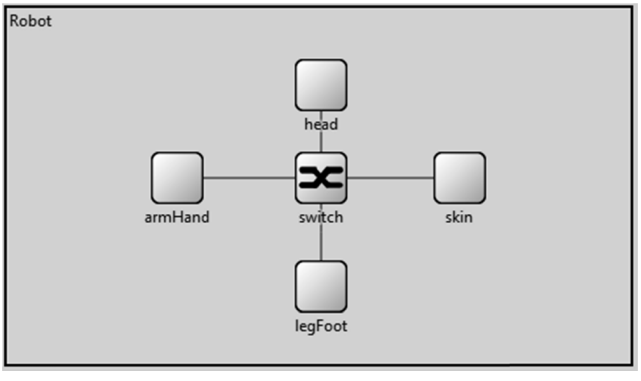

FIGURE 4. Star topology of the IRN Architecture simulator.

We tested the average end-to-end propagation delay in three different network architecture simulations under the condition of a full network load. The end-to-end propagation delay we tested is divided into two types, where one is the

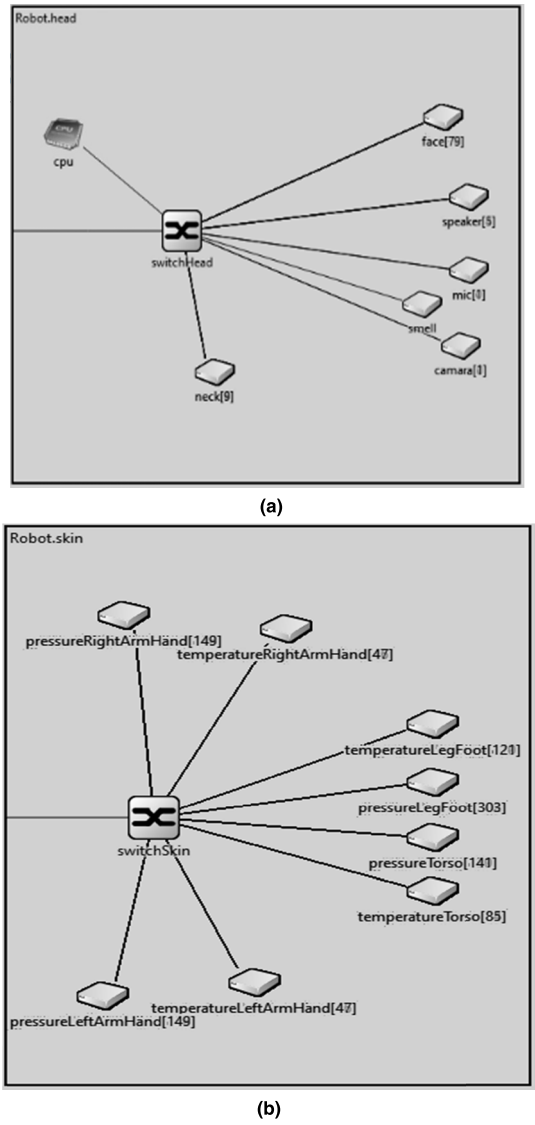

FIGURE 5. (a) Head domain and (b) Skin domain of the IRN Architecture simulator.

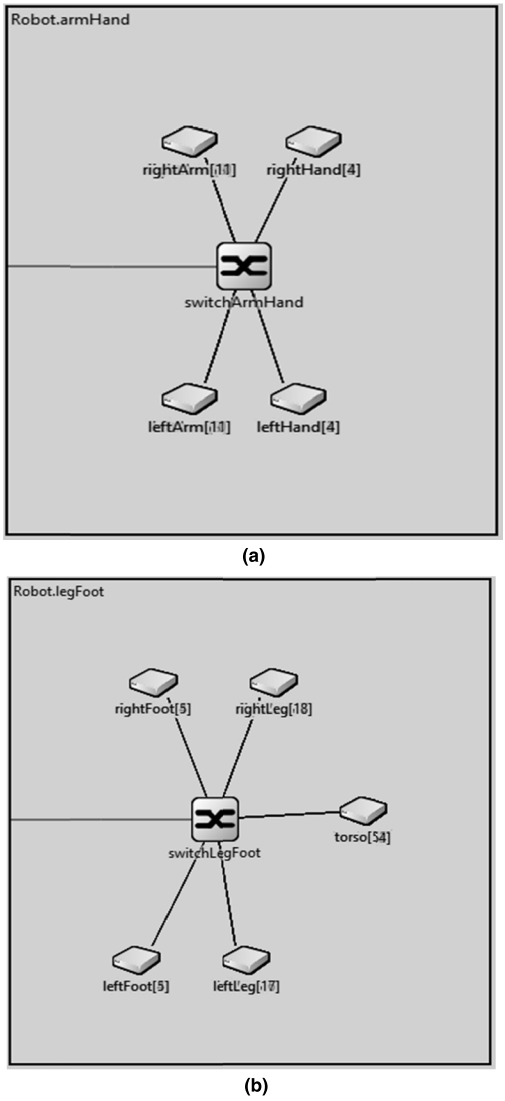

FIGURE 6. (a) Arm and hand domain and (b) Leg and foot domain of the IRN Architecture simulator.

delay generated when the information collected by the sensor is sent to the CPU. When sending data packets from the sensor to the CPU, we check the corresponding timestamps when the data are sent to the network link and when the data are received. The average propagation delay result can be calculated through these. Equation 7 shows the definition of the average end-to-end propagation delay,

$$D_p^t = \frac{\sum_{m=0}^{N_t-1} [T_r^t(m) - T_s^t(m)]}{N_t}, \tag{7}$$

where $T_s^t(m)$ is the timestamp reading which indicates the timing of data packets sent from the mth temperature sensor. and $T_r^t(m)$ is the timestamp result that indicates the timing of data packets sent from the mth temperature sensor received by the CPU. The range of m is from 0 to $N_t - 1$. Similarly, the average end-to-end propagation delay of the pressure sensor to the CPU can be obtained by the following equation:

$$D_p^p = \frac{\sum_{n=0}^{N_p-1} [T_r^p(n) - T_s^p(n)]}{N_p}, \tag{8}$$

where $T_s^p(n)$ is the timestamp reading at the time when data packets are sent from the nth pressure sensor. $T_r^p(n)$ is the timestamp reading that indicates the time when data packets sent from the nth pressure sensor are received by the CPU. The range of n is from 0 to $N_p - 1$.

The other end-to-end propagation delay is generated when the central processing unit sends a command to the actuator. The average end-to-end delay can be obtained in a similar way as shown in equation 9.

$$D_p^a = \frac{\sum_{k=0}^{N_a-1} [T_r^a(k) - T_s^a(k)]}{N_a}, \tag{9}$$

where $T_s^a(k)$ represents the timestamp reading of actuator control data sent by the CPU to the kth actuator. $T_r^a(k)$ is the timestamp reading that indicates the time when actuator control data packets are received by the kth actuator. In addition, N_a is the total number of actuators distributed in the head domain, arm and hand domain, and leg and foot domain.

$$N_a = N_a^H + N_a^{AH} + N_a^{LF}, \tag{10}$$

We have established two kinds of communication scenarios. In the first scenario, we tested the end-to-end delay performance of the three network architectures when the humanoid is in different environments and the sensor data changes to different degrees due to changes in environmental parameters. First, the network remains fully loaded, and the size of the data packet generated by the sensor is stable. This scenario means that the temperature and pressure of the humanoid's environment have not changed significantly. In this scenario, we set the median value of the temperature data generated by all temperature sensors to 25 degrees Celsius, and the amount of random fluctuation is controlled within ± 0.4 degrees Celsius. The median value of the random number generated by the pressure sensor is set to 1 kPa, and the amount of change is controlled under ± 10 Pa. Subsequently, in order to reflect the impact of sudden changes in surrounding environmental factors on the communication volume of the sensor network. We do not change the median value of the random number generated by all sensors but increase the range of change for some sensor values to 5 times the original value, with the intention of restoring a large data change in some sensors. We define that when the difference between the data value generated by the temperature or pressure sensor and the average value in a period of time reaches 1 degree Celsius or 50 Pa, the values of the two kinds of sensors reach the excited threshold and the related sensors are in the excited state, and vice versa call it stable state. In excited state, sensor data management will have to forward the data collected by these sensors to the CPU that acts the brain functionality. We first adjust the proportion of sensors in the excited state to 20% of the total number of sensors. Therefore, since the value change generated by these 20% of the sensors exceeds the threshold of sensor data management filtering data, it will generate a greater communication volume than the first solution. Then, we adjusted the sensor values of this solution to 40%, 60%, 80%, and 100% to compare the impacts of varying degrees of change in environmental factors on the average end-to-end delay of the entire network.

Figure 7 shows the average sensor to CPU end-to-end propagation delay results collected during the communication process of the three network architectures under different conditions of varying degrees of environmental factors. It can be found that since sensor data management is not used to filter the generated signals, the end-to-end propagation delay results of architecture 1 and architecture 2 in different environments are basically stable. The end-to-end propagation delay of architecture 1 has been maintained at approximately

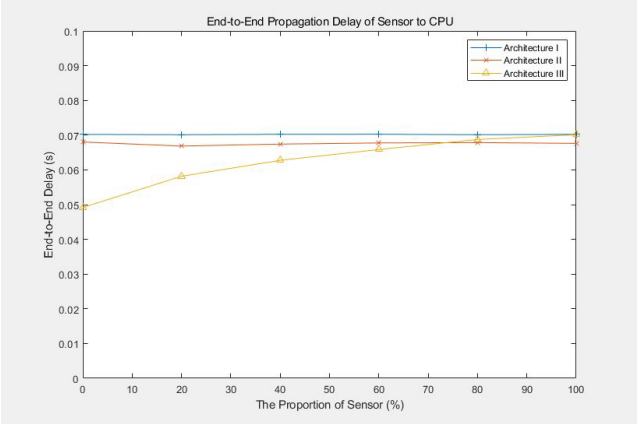

FIGURE 7. End-to-End propagation delay comparison of sensor to CPU communication.

70.5 ms, and since architecture 2 contains more switches than architecture 1, which reduces the congestion of the network, the end-to-end propagation delay is stable at approximately 67.6 ms. The results of architecture 3 show that the end-to-end propagation delay is 49.3 ms in a stable state where the environmental factor change is 0, which is quite different from the results for the other two architectures. The result of architecture 3 tends to move closer to the result of architecture 1 as the degree of change in environmental factors increases. When the proportion of sensors whose data change exceeds the threshold is 100%, the result is basically the same as that for architecture 1.

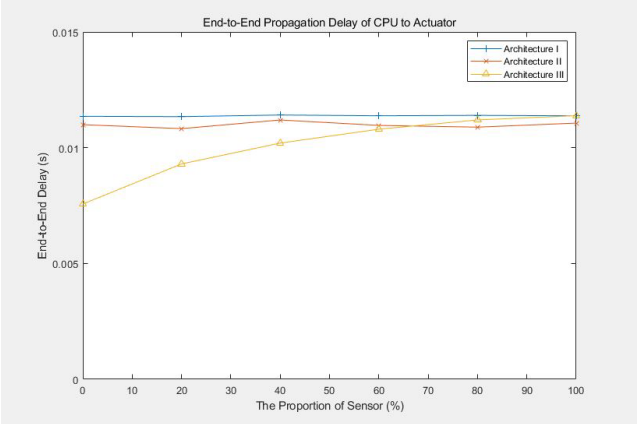

FIGURE 8. End-to-End propagation delay comparison of CPU to actuator communication.

Figure 8 shows the average CPU to actuator end-to-end propagation delay results of the three network architectures when the degree of environmental factor change is different. Similar to the sensor to CPU end-to-end, the delays in the results of architecture 1 and architecture 2 remain stable under different conditions, at approximately 11.7 ms and 11.0 ms, respectively. The result for architecture 3 increases with the change in the proportion of sensors whose data changes exceed the threshold, from 7.5 ms to 12.5 ms.

In the second scenario, we tested the average end-to-end delay changes generated by the three network architectures when reducing the bandwidth of the backbone network to compare the performances of the three architectures. In this scenario, the initial bandwidth of the backbone given to the three network architectures is consistent with scenario 1. At the same time, in order to compare the impacts of the number of sensors in the excited state on the network performance in this scenario, we conducted two sets of tests. The first test set keeps the data value content in the stable state. The other test set assumes that all the sensors are in the excited state. Based on the required backbone bandwidth, we reduced the bandwidth by 5% increments and performed several tests until the bandwidth was reduced to 70%.

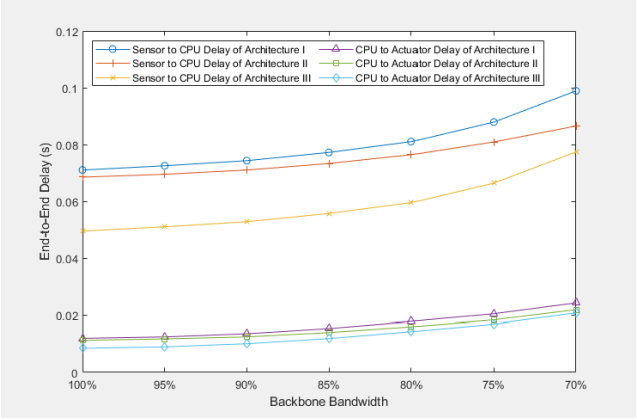

FIGURE 9. The average End-to-End delay of sensor to CPU and CPU to actuator changes with the reduction of backbone bandwidth under stable conditions of sensor-generated data.

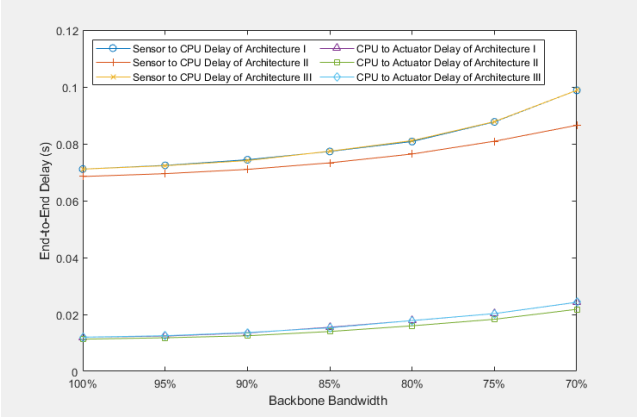

FIGURE 10. The average End-to-End delay of sensor to CPU and CPU to actuator changes with the reduction of backbone bandwidth under the condition of large changes in the data generated by all sensors.

In Figure 9, all the sensors of three architectures are in the stable state. On the other hand, all the sensors are in the excited state in Figure 10. This is to compare the performances in the two extreme cases. In architecture 3, compared to the other two architectures, due to the use of sensor data management, both the sensor-to-CPU and CPU-to-actuator delay are low. Because architectures 1 and 3 have the same network structure except for the sensor data management, as the bandwidth decreases, the delays show similar rising curves. Compared with architecture 2, the other two architectures show a larger upward trend when the bandwidth is reduced to 80%. In Figure 10, architectures 1 and 3 both start from a similar starting point, and the end-to-end delays show similar rising curves as the bandwidth decreases. Due to the multiple switches are used in architecture 2, the test result shows a smaller rising trend than the other two architectures.

V. CONCLUSION

In this paper, efficient IRN architectures for humanoid robots are proposed and analyzed with simulations. To achieve equivalent human perception and motor capabilities, large numbers of sensors and actuators need to be implemented in a humanoid robot. To maintain good data transmission conditions between sensors, actuators and processors, a complex and reliable network architecture is essential. Considering the differences in physical locations and functions of sensors and actuators as the starting point, the IRN is first divided into several domains. Each domain contains a large number of network nodes and switches with similar functionalities and capabilities. The switches connect the nodes to the backbone network. In addition, according to the differences in the bandwidth requirements of the different domains, 10 Mbps, 200 Mbps and 10 Gbps Ethernet network links are assumed.

In addition, this paper proposes three IRN network architectures and verifies their communication delays in several cases including fully loaded network environments. The results show that in such environments, although the three architectures can maintain the delay within an acceptable range when sufficient bandwidth is allocated, due to the differences in network architecture design, there are significant differences in the test results of the three architectures under different communication conditions. Therefore, from a biological point of view, these delays obtained through simulation are similar to the propagation speed of electrical signals in the human nervous system.

In architecture 2, the number of switches is greater than that in the other architectures. Although the increase in the number of switches would increase the physical weight of the network components to a certain extent, the network congestion may be reduced accordingly.

In architecture 3, the sensor data management device may effectively reduce the repeated transmission of redundant signals sent by the sensor in a stable environment. The reduction in the amount of data brought about by the filtering of either repetitive or nonessential data may greatly reduce the transmission delay and the backbone network traffics.

In future research, a set of new network communication protocols and standards that cope with the network architecture requirements may be studied. For example, network protocols may address the physical weights of communication devices and traffic engineering such as priorities, credit-based

shaping, bypassing, and preemption. Additionally, partial failures and survivability of networks should be considered.

REFERENCES

[1] L. Chassagne, O. Bruneau, A. Bialek, C. Falguiere, E. Broussard, and O. Barrois, "Ultrasonic sensor triangulation for accurate 3D relative positioning of humanoid robot feet," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2856–2865, May 2015.

[2] T. Kim and S. Park, "Equivalent data information of sensory and motor signals in the human body," *IEEE Access*, vol. 8, pp. 69661–69670, Apr. 2020.

[3] H.-P. Huang, J.-L. Yan, T.-H. Huang, and M.-B. Huang, "IoT-based networking for humanoid robots," *J. Chin. Inst. Engineers*, vol. 40, no. 7, pp. 603–613, Oct. 2017.

[4] K. W. Chung and H. M. Chung, Eds., *Gross Anatomy Board Review*. Hagerstown, MD, USA: Lippincott Williams, 2005, p. 364.

[5] Y. J. Li, W. C. Chou, C. Y. Chen, B. Y. Shih, L. T. Chen, and P. Y. Chung, "The development on obstacle avoidance design for a humanoid robot based on four ultrasonic sensors for the learning behavior and performance," presented at the Int Conf. Ind. Eng. Eng. Manage., Macao, China, Dec. 7-10, 2010.

[6] V. Favot, T. Buschmann, M. Schwienbacher, A. Ewald, and H. Ulbrich, "The sensor-controller network of the humanoid robot LOLA," presented at the Int. Conf. Humanoid Robots, Osaka, Japan., Nov. 2012.

[7] F. Kanehiro, Y. Ishiwata, H. Saito, K. Akachi, G. Miyamori, T. Isozumi, K. Kaneko, and H. Hirukawa, "Distributed control system of humanoid robots based on real-time Ethernet," presented at the Int Conf. Intell. Robots Syst., Beijing, China, Oct. 2006.

[8] M. Ramanathan, N. Mishra, and N. M. Thalmann, "Nadine humanoid social robotics platform," presented at the Int. Conf. Comput. Graph., Calgary, AB, Canada, Jun. 2019.

[9] J. Mossbridge and E. Monroe, "Team Hanson-Lia-SingularityNet: Deep-learning assessment of emotional dynamics predicts self-transcendent feelings during constrained brief interactions with emotionally responsive AI embedded in Android technology," unpublished.

[10] A. Goswami and P. Vadakkepat, "ASIMO and humanoid robot research at honda," in *Humanoid Robotics: A Reference*. Dordrecht, The Netherlands: Springer, 2019, pp. 55–90.

[11] S. Kuindersma, R. Deits, M. Fallon, A. Valenzuela, H. Dai, F. Permenter, T. Koolen, P. Marion, and R. Tedrake, "Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot," *Auto. Robots*, vol. 40, no. 3, pp. 429–455, Jul. 2015.

[12] J. Lee and S. Park, "Time-sensitive network (TSN) experiment in sensor-based integrated environment for autonomous driving," *Sensors*, vol. 19, no. 5, p. 1111, Mar. 2019.

[13] *LIN Specification Package—Revision 2.2A*, Standard ISO 17987 Part 1-7, 2010.

[14] G. Cena and A. Valenzano, "FastCAN: A high-performance enhanced CAN-like network," *IEEE Trans. Ind. Electron.*, vol. 47, no. 4, pp. 951–963, Aug. 2000.

[15] D. Paret, *Multiplexed Networks for Embedded Systems: CAN, LIN, FlexRay, Safe-by-Wire*. Chichester, U.K.: Wiley, 2007.

[16] N. Navet, Y. Song, F. Simonot-Lion, and C. Wilwert, "Trends in automotive communication systems," *Proc. IEEE*, vol. 93, no. 6, pp. 1204–1223, Jun. 2005.

[17] R. Makowitz, "FlexRay-a communication network for automotive control systems," in *Proc. IEEE Int. Workshop Factory Commun. Syst.*, Jun. 2006, pp. 207–212.

[18] *IEEE Draft Standard for Ethernet Amendment 4: Physical Layer Specifications and Management Parameters for 10 Mb/s Operation and Associated Power Delivery Over a Single Balanced Pair of Conductors*, Standard 35.100.10 35.110, Physical layer Networking, Standard P802.3cg/D3.1, Jun. 2019.

[19] (2019). *IEEE P802.3ck Task Force-Tools and Channels*. [Online]. Available: https://www.ieee802.org/3/ck/public/tools/index.html

[20] *IEEE Draft Standard for Ethernet Physical Layer Specifications and Management Parameters for Greater Than 1 Gb/s Automotive Ethernet*, Standard 35.100.10 35.110, Physical layer Networking, Standard P802.3ch/D3.2, Feb. 2020. 2020.

CHENGYU CUI received the B.S. degree from the Beijing University of Technology, Beijing, China, in 2014. He is currently pursuing the combined M.D.-Ph.D. degree with the Department of Electronic Engineering, Hanyang University, Seoul, South Korea. His research interests include in-vehicle networks, in-robot networks, the Internet of Things (IoT), and cable broadcasting.

SUNGKWON PARK (Senior Member, IEEE) received the B.S. degree from Hanyang University, Seoul, South Korea, in 1982, the M.S. degree from the Stevens Institute of Technology, Hoboken, NJ, USA, in 1983, and the Ph.D. degree from the Rensselaer Polytechnic Institute (RPI), Troy, NY, USA, in 1987, all in electrical engineering. He joined Tennessee Technological University, Cookeville, TN, USA, as an Assistant Professor with the Department of Electrical Engineering, upon his graduation from RPI, in 1987. He left Tennessee Technological University, in 1993, as a Tenured Associate Professor to join Hanyang University. He is currently a Full Professor with the Department of Electronic Engineering, and the Dean of the College of Engineering, Hanyang University. He has published nine authored books and over 200 refereed technical articles in the areas of CATV, IPTV multimedia systems, and vehicle communication systems. He holds over 50 domestic and international patents. His research interests include in-vehicle networks, in-robot networks, and autonomous vehicles. He is also a Lifelong Member of the Korean Information and Communications Society. He received the Initiation Award from the U.S. National Science Foundation, in 1990. He has served as the Chairperson for the Digital Cable TV Steering Committee representing the Ministry of Information and Communications (MIC), from 2001 to 2005. He has been serving as the Chairperson for the Korea Digital Cable Forum, since April 2008.