Journal of Innovative Technology, Vol. 3, No. 2, pp. 51-56, September 2021
http://doi.org/10.29424/JIT.202109_3(2).0006

Hand Gesture-Recognition-Based Home Automation

Shun-Jen Hsiao [1], Kuang-Hui Chi [2*], and Chian C. Ho [3]

ABSTRACT

This study covers a design and implementation of a hand gesture-recognition capable device that can be geared to an IoT (Internet of Things) environment with household security, remote power control of appliance, indoor temperature monitoring, user's heart rate sensing and location positioning. Interactivity is realized through contactless hand gesture recognition whereby control over electrical apparatus is accomplished via simple hand signs and swipe for ON/OFF manipulation. Such operation modus is fitting for aged individuals who have ataxia or muscle incoordination. Activated by a server with value-added functions, all the intended manipulations can be invoked with a WWW browser or an Android APP. Our end user interface is targeting a wearable device with off-the-shelf wireless technology interconnecting legacy appliances for increased automation. Experiments show that our implementation achieves a successful recognition rate of over 96% on average. Hand gestures indicative of commands for different types of apparatus result in a promising IoT application that lends elegantly itself to a variety of fields in intelligent living space.

Keywords: Home automation, Internet of Things, wireless network, protocol, wearable.

1. INTRODUCTION

This research is aimed at mitigating inconvenience of home appliances in daily life. Such inconvenience arises, for example, when a remote paired with a certain equipment like a TV set or an air conditioner is forgotten or cannot be found occasionally. Other common occurrences are that, for turning on/off electrical items like lamps and electric fans, the user is often required to approach within the reach of electrical sockets or wall switches, bringing certain hardship to physically handicapped seniors. To date, many kinds of electrical or electronic products have been designed to operate independently with no means of unifying their connectivity.

To remedy this lack of connectivity, if household electrical devices can be subsumed into residential premises to a fuller extent, home automation could be made possible through existing wireless technologies such as IEEE 802.11 Wi-Fi, BLE (Bluetooth Low Energy), or 6LoWPAN (IPv6 over Low-Power Wireless Personal Area Networks). These have all been used successfully to transmit user commands over the air to control appliances (Hsieh *et al*. 2014).

This paper examines a prototype device that could be reduced to a wearable size, which captures user commands for different electrical devices. Unlike today's touchscreen-oriented products, this prototype does not require fingertip contact or fine finger gestures that would be a challenge for the elderly. A contactless mechanism was deemed more appropriate for older users or those with poor sight and reduced movement ability. Touchscreens are also known to pose issues to users that have uncontrollable muscle tremors.

Currently, mainstream wearable technology and IoT (Internet of Things) applications have been focused on health management (Mano *et al*. 2016; Santos *et al*. 2014). This means that a unified application programming interface (API) is still not unavailable. Having a mixture of proprietary specifications to coordinate raises numerous interoperability issues. In addition to this, there are legacy household appliances like kettles and lamps that are standalone, with little to no automation. These issues are addressed by the introduction of a translation device that mediates between devices for increased convenience. The translation device enables wireless functionality for control over previously inert interfaces such as AC (alternating current) sockets. Inspired by a health care approach, this prototype wearable is also equipped with a heart rate monitor, a video surveillance camera, and a global positioning system (GPS) module that would hopefully be useful to emergency responders in the event of an accident or illness. The devices could thus be used to contact family or caregivers if the wearer's physiological situation were to change suddenly. Location information is essential for the elderly to prevent accidents.

There is considerable literature available on human-computer interaction. This research project was driven by a number of surveys (Oudah *et al*. 2020; Yasen and Jusoh 2019) which address the adaptation of technology for less physically able individuals, or individuals with mental impairment. Many studies focused on a glove with sensors that detects finger and hand movements (Králik

Manuscript received August 24, 2021; accepted September 9, 2021.
[1] Assistant Researcher, Department of Electrical Engineering, National Yunlin University of Science and Technology, Taiwan 640301, R.O.C.
[2*] Professor (corresponding author), Department of Electrical Engineering, National Yunlin University of Science and Technology, Taiwan 640301, R.O.C. (e-mail: chikh@yuntech.edu.tw).
[3] Associate Professor, Department of Electrical Engineering, National Yunlin University of Science and Technology, Taiwan 640301, R.O.C.

and Šuppa 2021). This echoed a growing research trend where machine learning is used to adapt technology to a user (Zhang *et al*. 2019; Nguyen *et al*. 2021). The prototype used in this paper distinguishes itself from previous work in three significant ways:

- Our design and implementation are lightweight that is runnable on a low-cost platform like Raspberry Pi with limited computing power yet maintaining performance. Hence, this research reflects cost-effectiveness of adopting a simple single-board platform.
- More than hand gesture recognition, we provide an avenue to interrelate home apparatus under a unified user interface to materialize a prototype IoT application. Accordingly, our design tenet lends itself to other applications in intelligent living space.
- Our image processing used is complimentary to but not competing with machine learning-based schemes subject to the limited amount of training data. As evidenced in (Shih and Chi, 2019), machine learning may underperform if the amount of training data is limited; joint use of image processing techniques can strengthen the quality of training data and thereby greatly benefit classification tasks.

The remainder of this paper is organized as follows: Section 2 outlines the development of the prototype, Section 3 describes the hand gesture-recognition mechanism; experimental results are shown in Section 4, and Section 5 concludes the study.

2. ARCHITECTURE

Figure 1 shows the overall architecture which is comprised of two primary parts: (1) a wearable/portable device, and (2) a home IoT ecosystem. The former is managed by a Raspi and a small video camera (Pi NoIR), a heart rate sensor unit, and a mounted GPS module, shown in Fig. 2. The latter is the home IoT environment that involves the establishment of a wireless network embodied with another Raspberry Pi, a base communicating with other household modules including a door lock that is used for security, and remote-control access for two AC sockets, an IR (infrared) remote for the air conditioner, and indoor temperature sensing (Fig. 3.) Concerning home security, when a user wishes to enter or leave the house, a predefined hand gesture indicating "lock" or "unlock" the door suffices. Alternatively, the locking and unlocking can be performed with the app developed to serve this ecosystem.

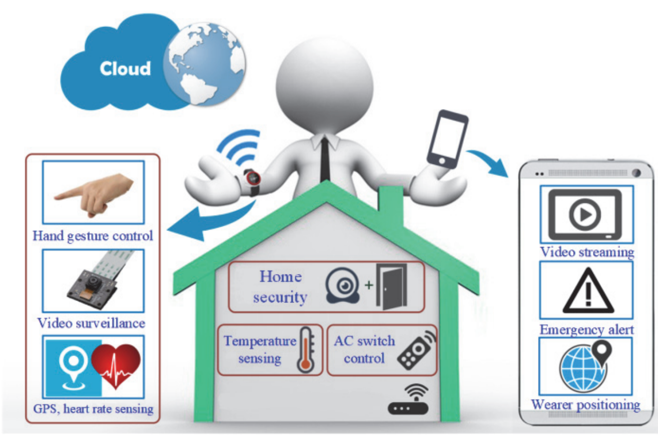

Fig. 1 Overall architecture.

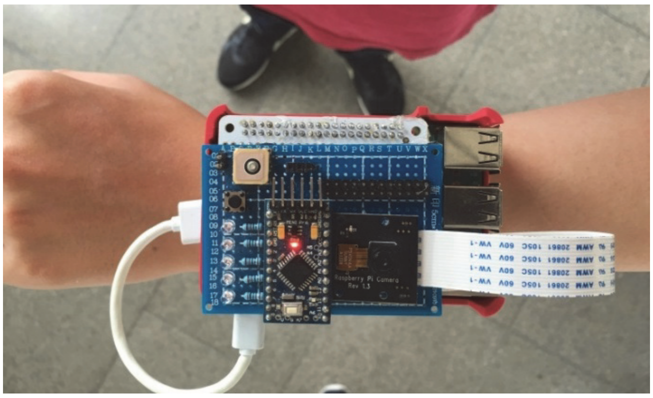

Fig. 2 A prototype of the end user interface.

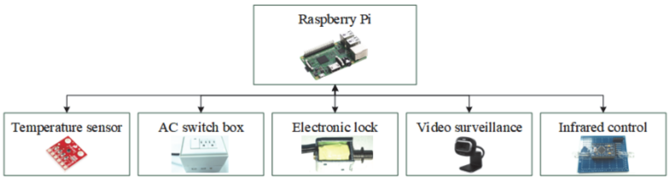

Fig. 3 Our home IoT is materialized with a residential control center communicating with five different modules over the wireless medium.

Our home IoT ecosystem includes an indoor temperature sensor and is also able to control the AC power to electrical appliances. For instance, if the indoor temperature becomes high/low, the user is enabled to turn on/off an electrical fan or an air conditioner by associated gestures plus hand waving. In case the temperature continues rising, there are probably fire incidents in proximity and alerts should be raised, so as to prevent possible loss of properties. For more practical uses, a web server plus a MySQL database system were added to record and respond to requests for domestic status tracking and/or live video streaming for surveillance purposes. All the responded data are accessible from a smartphone over the Internet anywhere anytime. In this manner, the server is tasked to notify the house owner or concerned family members instantly of any potential danger is about to occur.

Our home IoT takes shape by extending a platform, Raspberry Pi, with feature functionalities as shown in Fig. 3. All the functionalities are incorporated through WiFi interfacing into the platform, namely the residential control center. Feature functionalities are hardware-based but software-triggered at user's commands. User commands originate from the hand gesture system or from manual operations on a web-based graphic user interface (GUI) (Fig. 4). Among feature functionalities, we highlight the remotely controllable AC sockets and an IR remote controls for an air conditioner. Such AC sockets are soldered in a standard outlet box of 12.5 × 7.5 × 5 cm in size (Fig. 5) that contains circuits with solid-state relays keeping the power at each socket on or off under user's directives. The box also embeds circuits converting power from AC 110V to DC (direct current) 5V, whose outputs act as two 5V DC power supplies to other electronics appliances.

Many consumer electrical and electronic products are paired with IR remotes. For compatibility, this home IoT is aligned with such commodities yet free from changing their machinery; we provide a hardware-software codesign for common household appliances to the user's benefits. While each device was originally

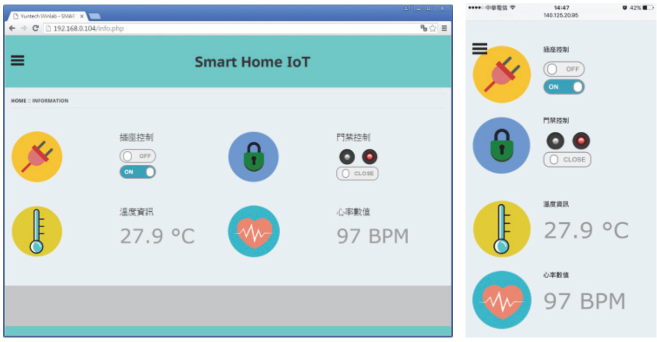

Fig. 4 Screenshots of our webpage and APP.

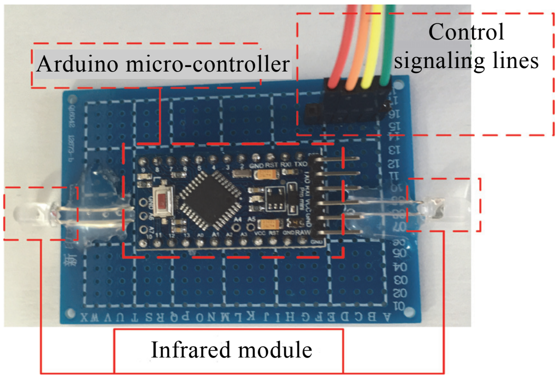

Fig. 6 IR remote for an air conditioner.

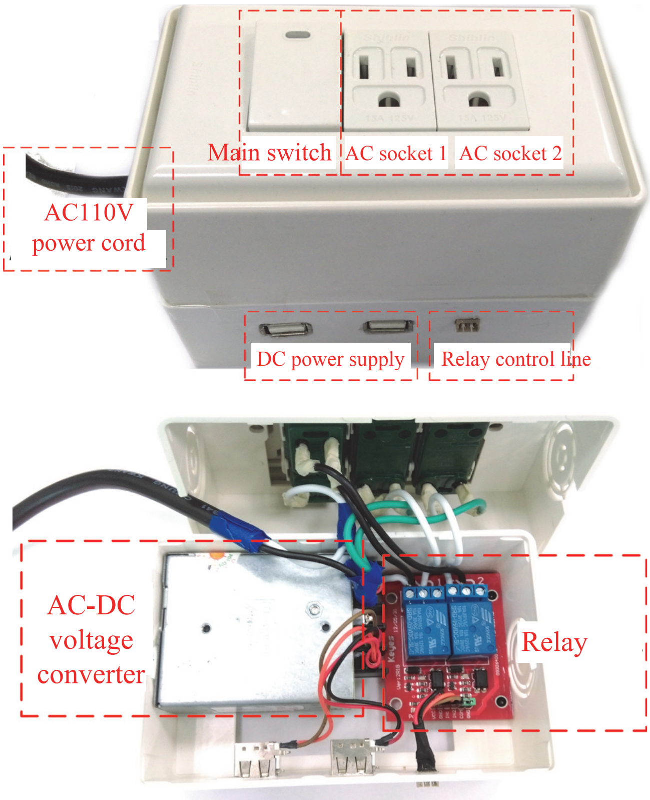

Fig. 5 A remotely controllable box of AC sockets.

meant to work with a specific remote, there are a plurality of remotes around, causing them likely to be misplaced. As a remedy, we unify various remotes into a single interface by letting the wearable identify hand gestures and communicate, possibly through the Internet, with the residential control center that generates necessary commands over infrared to the controlled devices. Accordingly, our wearable has hyper control over a variety of apparatus. In the system, IR control is implemented over Arduino (Fig. 6) that transmits manufacturer-defined IR codes under the charge of the residential control center. The Arduino widget can be used in different rooms inside a house, maintaining wireless connectivity with the residential control center over WiFi, BLE, or ZigBee, with a saving of deployment costs. The distributed widgets operate in place of original remotes that control home appliances within range with increased accessibility.

3. HAND-GESTURE RECOGNITION

Let us now proceed to elaborate on how we carry out hand gesture recognition with image processing techniques. Our current design emphasizes rapid yet stable recognition under a less complex background and under sufficient light conditions, as clarified below.

3.1 Development Environment

Our wearable device was prototyped chiefly by using a Raspberry Pi and a NoIR video camera module to implement hand gesture recognition. The recognition undergoes four stages: image acquisition, image preprocessing, region identification, and feature resolution. Note that the output of the last stage is fed to the residential control center over the wireless medium for subsequent appliance manipulation.

3.2 Hand Sign

In order to make the gestures as intuitive as possible, finger-counting and mid-air hand swiping were used. The upward-facing Pi NoIR camera on the wearable captures hand images in motion. As shown in Fig. 7, gestures of finger numerals 1 to 5 indicate different directives for control over the first AC socket, the second AC socket, the door lock, playing music, and the air conditioner, respectively. Gesture recognition begins when the user poses his/her hand still with at least one finger raised for 3 seconds within detected range. As soon as a certain hand gesture has been identified, the associated device is brought under immediate control of the user. The user can then swipe left (right) to turn the associated device on (off) or swipe left to open and right to close. So far the system is only programmed to respond to left/right swipes; however, a wider array of instructions is possible.

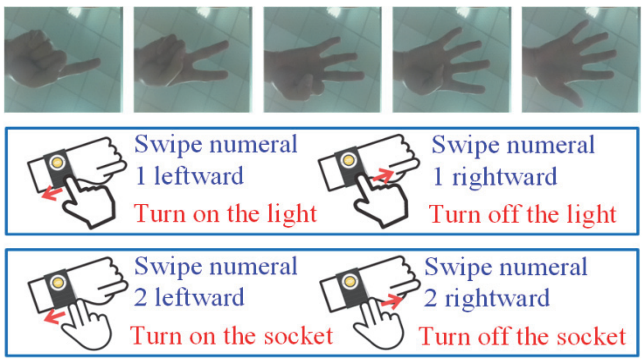

Fig. 7 Hand gesture and swipe above the wearable device for directives signaling

3.3 Methodology

Figure 8 maps the image recognition process that was developed using OpenCV (Bradski and Kaehler 2017). In order to highlight characteristics for these hand gestures, certain parts of the image set were filtered out to emphasize the hand. A convex-hull search was used to determine the contours of the hand. This contour was also important to resolve the parameters for user instructions. In computational geometry, a convex hull can be applied to find concave areas of an object. In this case, the concavities of a hand gesture are needed. A sample of a convex hull enclosing a hand gesture, with concavities is shown in the bottom-right corner of Fig. 8.

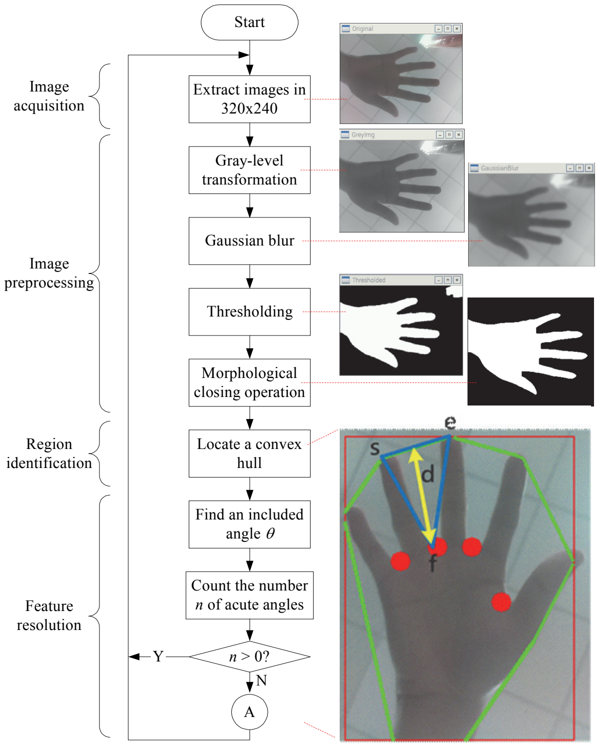

Fig. 8 Flowchart of our gesture recognition process. For better visualization, images resulting from respective steps are illustrated to the right alongside.

When one convex hull has been identified, the system continues searching for more. There may be several concave areas, formed by consecutive fingertips or a fork between a pair of fingers. In this situation, the system detects a *concavity* based on 4 parameters: two adjacent fingertips *s* and *e*, the fork *f* where the ends of two fingers meet, and the distance *d* from *f* to the edge *se* of the convex hull (see an illustration on the bottom-right quadrant of Fig. 8.) Since different gestures produce a varying number of concavities, gestures can be determined by completing steps 2 to 5 below:

Step 1: Calculate the lengths of sides $\overline{se}$, $\overline{sf}$, and $\overline{ef}$. Where the coordinates of these points are $s(sx, sy)$, $e(ex, ey)$, and $f(fx, fy)$, respectively, we have

$$\overline{se} = \sqrt{(s_x - e_x)^2 + (s_y - e_y)^2};$$

$$\overline{sf} = \sqrt{(s_x - f_x)^2 + (s_y - f_y)^2};$$

$$\overline{ef} = \sqrt{(e_x - f_x)^2 + (e_y - f_y)^2}.$$

Step 2: Find the included angle θ between two adjacent fingertips by the law of cosines, *i.e.*,

$$\theta = arccos\left(\frac{\overline{sf}^2 + \overline{ef}^2 - \overline{se}^2}{2 \cdot \overline{sf} \cdot \overline{ef}}\right)$$

Step 3: Let n count concavities, each whose included angle θ is less than 90 degrees, *i.e.*, *n* is the number of acute angles. Such counting is based on a fact that any two adjacent fingers of a man's open hand typically cannot form an included angle over 90°. The variable *n* is well reflective of what gesture we have (hand sign with numeral *n* + 1). For instance, if *n* equals 1, the user is making a hand gesture with 2 fingers raised.

However, in event that *n* is 0, further actions determined fingertips and palmar parameters to resolve whether a hand sign with numeral 1 is made, as described in the following steps.

Step 4: Estimate the centroid of the palm by computing the image moment. An image moment allows for pixel intensities $I(x, y)$ for any pair of coordinates (x, y). Letting $g(x, y)$ represent a digital image, for the 2D continuous function $g(x, y)$ the raw moment of order $(p + q)$ is defined as

$$M_{pq} = \sum_x \sum_y x^p y^q I(x, y).$$

In addition, $M_{00} = \sum_x \sum_y I(x, y),$

$$M_{10} = \sum_x \sum_y x \cdot I(x, y),$$

$$M_{01} = \sum_x \sum_y y \cdot I(x, y)$$

The centroid is at $(\tilde{x}, \tilde{y}) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}}\right)$, as illustrated in the right column of Fig. 9.

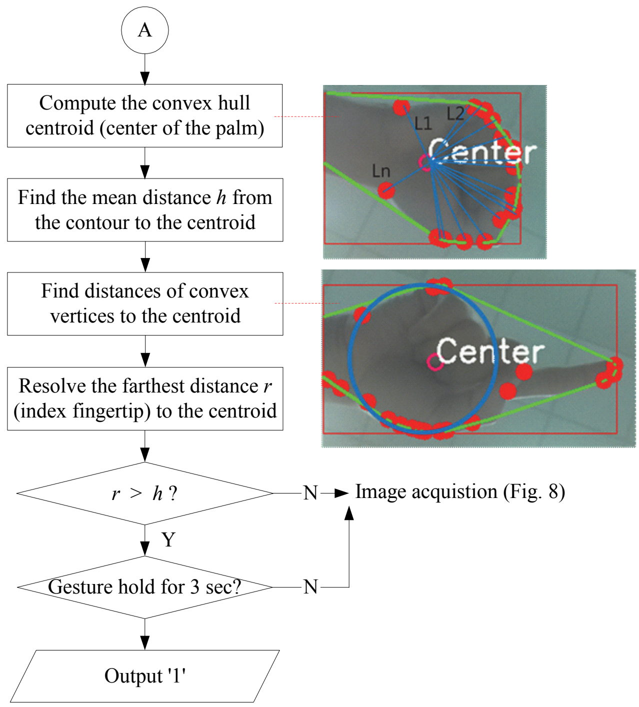

Fig. 9 Flowchart of determining finger numeral 1.

When hand gestures alter, changes of the centroid are recorded accordingly. If the gesture shifts in position (due to hand swiping), location changes of the centroid indicate the direction of motion because of hand swiping left or right.

Step 5: Calculate the mean distance from the contour of the palm to the centroid (excluding fingers.) First locate the forks of the fingers through a convex-hull, and search to determine points on that contour that are inside the convex envelope. For each point found k, the distance to the centroid, L_k to k is calculated and thus the arithmetic mean can be calculated giving the radius of the palm.

$$\left(\sum_{k=1}^{n} L_k\right) / n$$

Step 6: Find the distances of convex-hull vertices to the centroid. As the convex envelope encloses the contour of the hand, it is possible to calculate the distances of convex-hull vertices to the centroid of the palm. When a vertex lies evidently farther than other ones, we deduce that the distant vertex indicates the index fingertip. For example, provided that {150, 40, 50, 20} is a set of such distances in question, the element 150 stands out from the set, implying that 150 is the distance of the tip of the index finger to the center of the palm.

4. EXPERIMENTAL RESULTS

A series of experiments were conducted to validate the performance of the design with regard to recognition rate and speed. Experiments were undertaken by letting our wearable device carry out recognition of 10 testers making hand gestures. To strengthen our results while maintaining a fair basis for comparison, the wearable was placed still facing up on a table for preventing interference from wrist shaking (Fig. 10.) All the testers were male who were invited to make gestures in order of numerals 1 to 5 prescribed in our use cases. Our residential control center was run at the backend to enable home IoT operations jointly.

Each gesture (finger numeral representation with hand swiping) was repeated 20 times per tester. So, a total of 200 samples were collected for every hand sign. Whenever a hand gesture was recognized, its connected appliance needed to be activated or deactivated to confirm recognition success or failure. Table 1 lists experimental results of different gestures, showing that our system achieves a successful recognition rate of over 96% on average. Even though identifying the hand gesture of numeral 1 is slightly inferior, it does not cause performance discrepancy.

Concerning operational timeliness, we measured how much time was spent recognizing each hand sign correctly. Figure 11 reveals that recognitions were completed within 47 to 69 ms. Observe that, similar to traits of numeral 1 reflected in Table 1, the process of recognizing the hand gesture of numeral 1 took longer than that of other gestures for involving three more steps (Steps 4 to 6 of Section III.C.) Such time delay appeared so insignificant that a user could hardly perceive timing differences. This corroborates the real-timeliness of our system in practice.

To conclude this section, we remark on conditions that delimit the effectiveness of this study. Since our approach is vision-based, there are scenarios where the recognition process may fail because the acquired image does not contain a complete hand shape. As exemplified in Fig. 12, the left subfigure results from the user placing his hand too close to the camera, so the oversized hand gesture becomes unrecognizable. The right subfigure results from the subject moving out of the camera shooting range, so the hand gesture ends up becoming unidentifiable for our system. Another possibility arises when the user swipes hand gestures too fast to be captured distinctly. Other possible causes include slanting hand shape or images being shot in poor brightness. Any of afore-mentioned cases may fail the recognition process, leading to no response from the wearable device nor activity alteration from any of intended appliances. These conditions also explain why the recognition results of Table 1 cannot reach 100%. Such recognition issues can arguably be avoided by familiarizing the user with reasonable limitations of the device.

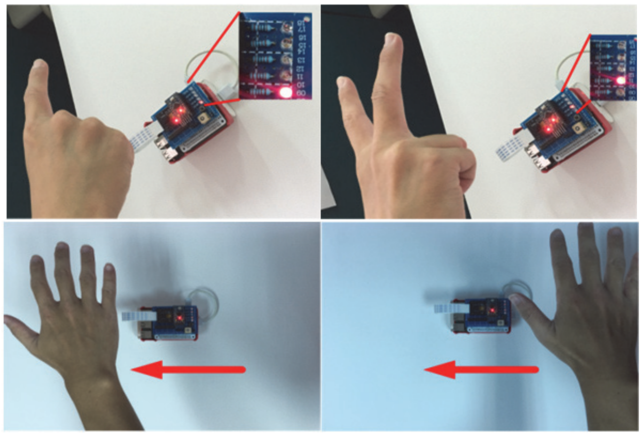

Fig. 10 Hand gesturing samples during experiments.

Table 1 Statistics of experimental results.

Tester \ Gesture	1	2	3	4	5
Subject 1	18/20	20/20	20/20	20/20	20/20
Subject 2	19/20	20/20	20/20	20/20	20/20
Subject 3	20/20	20/20	19/20	19/20	20/20
Subject 4	20/20	20/20	19/20	20/20	19/20
Subject 5	20/20	19/20	19/20	20/20	20/20
Subject 6	18/20	19/20	20/20	19/20	20/20
Subject 7	20/20	19/20	20/20	20/20	20/20
Subject 8	20/20	19/20	18/20	20/20	20/20
Subject 9	18/20	20/20	20/20	20/20	18/20
Subject 10	19/20	20/20	19/20	19/20	18/20
Recognition rate	192/200	196/200	194/200	197/200	195/200
	96%	98%	97%	98.5%	97.5%

```
68.657516 (ms) , Gesture(1) 48.437882 (ms) , Gesture(2) 47.285496 (ms) , Gesture(3)
68.798973 (ms) , Gesture(1) 48.437362 (ms) , Gesture(2) 48.250071 (ms) , Gesture(3)
69.019387 (ms) , Gesture(1) 48.434392 (ms) , Gesture(2) 48.368612 (ms) , Gesture(3)
68.597203 (ms) , Gesture(1) 48.952409 (ms) , Gesture(2) 47.458098 (ms) , Gesture(3)
69.061835 (ms) , Gesture(1) 49.170272 (ms) , Gesture(2) 48.491944 (ms) , Gesture(3)
51.955925 (ms) , Gesture(4) 52.630139 (ms) , Gesture(5)
51.863114 (ms) , Gesture(4) 52.891959 (ms) , Gesture(5)
51.351608 (ms) , Gesture(4) 53.336122 (ms) , Gesture(5)
51.459733 (ms) , Gesture(4) 52.327225 (ms) , Gesture(5)
51.906343 (ms) , Gesture(4) 52.452380 (ms) , Gesture(5)
```

Fig. 11 Processing time of each hand gesture.

Fig. 12 Example scenarios where hand gesture recognition is likely to fail.

5. CONCLUSION

This paper proposed a hardware-software co-design and implementation of a device that recognized hand gestures indicative of commands for different types of household equipment. Originally dispersed legacy appliances could be made better coupled that worked together in a unified way catering for the user needs. Intuitive hand signs suffice for our purpose. We took an approach that upheld user-friendliness with ease of operations. Our design benefited the elderly or individuals who had difficulty in using touchscreen products. Apart from hand gesture recognition, our developed wearable embedded sensing units for monitoring physiological dynamics and user location.

In our architecture, the wearable device acted as an enabler for IoT application. Working in synergy with the wearable was a backend platform (a Raspberry Pi running a web server plus a database system) in charge of value-added functions: locking/unlocking of a door, video surveillance, remote powering up/off AC of sockets for plugged-in home appliances, and temperature sensing. Experiments and performance discussion indicated high recognition rate, implying the usefulness of our development. Section IV also highlighted certain conditions impairing our design so that the reader can gain a more comprehensive understanding of this research.

This research was a feasibility study to test the interaction of IoT systems with legacy devices and gesture based systems. The current wearable device was prototyped on a Raspberry Pi, which was however not fully convenient to wear on wrist. One practical slim workaround is to adopt Raspberry Pi Zero W as the platform for the end user interface. Adopting the alternative with miniaturization of battery power supply shall bring forth a lightweight device. In the future, concerning care for the dementia elderly, the wearable device can be transformed to a chest badge that is extended to be 4G/5G capable and carry an embedded speaker. Like a body orientation, the chest badge is empowered, with joint use of GPS, to guide the wearer through voice to his home or to the nearest police station in case he loses direction in town. The wearer's family can also be notified in an automated manner and soothes the user in worries through voice, so as to prevent accidents.

REFERENCES

Bradski, G. and Kaehler, A. (2017). Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library, O'Reilly.

Chen, Y.T. and Tseng, K.T. (2007). "Multiple-angle hand gesture recognition by fusing SVM classifiers", In *Proc. 3rd IEEE Int'l Conf. Automation Sci. and Engr*., 527-530.

Enea, C., Charlotte, F., Bam, S.S., Gemma, T., Lyes, K., Payvand Melika, P., and Donati Elisa, D. (2020), "Hand-gesture recognition based on EMG and event-based camera sensor fusion: A benchmark in Neuromorphic Computing", *Frontiers in Neuroscience*, **14**, Article 637.

Hsieh, C.-W., Chi, K.-H., Jiang, J.-H., and Ho, C. C. (2014) "Adaptive Binding of Wireless Devices for Home Automation", *IEEE Wireless Commun*, **21**(5), 62-69.

Jing, L., Yamagishi, K., Wang, J., Zhou, Y., Huang, T., and Cheng, Z. (2011). "A unified method for multiple home appliances control through static finger gestures", In *Proc. IEEE/IPSJ Int'l Symp. Applications and the Internet*, Japan, 2011.

Králik, M. and Šuppa, M. (2021), "WaveGlove: Transformer-based hand gesture recognition using multiple inertial sensors" (to appear in) *Proc. 29th European Signal Processing Conf.*

Lu, Z. Chen, X., Li, Q., Zhang X., and Zhou, P. (2014). "A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices", *IEEE Trans. Human-Machine Syst*., **44**(2), 293-299.

Mano, L.Y., Faiçal, B.S., Nakamura, L.H.V., Gomes, P.H., Libralon, G. L., Meneguete, R. I., Filho, G. P. R., Giancristofaro, G. T., Pessin, G., Krishnamachari, B., and Ueyama, J. (2016). "Exploiting IoT technologies for enhancing health smart homes through patient identification and emotion recognition", *Comp. Commun*, **89-90**, 1-13.

Nguyen, N.-H., Phan, T.-D.-T., Kim, S.-H., Yang, H.-J., and Lee, G.-S. (2021), "3D Skeletal Joints-Based Hand Gesture Spotting and Classification", *Applied Sciences*, **11**(10), 4689; https://doi.org/10.3390/app11104689

Oudah, M., Al-Naji, A., and Chahl, J. (2020). "Hand gesture recognition based on computer vision: A review of techniques", *Journal of Imaging*, **6**(8), 73-101. https://doi.org/10.3390/jimaging6080073

Santos, A., Macedo, J., Costa, A., and Nicolau, M.J. (2014). "Internet of Things and smart objects for m-health monitoring and control", *Procedia Technology*, **16**, 1351-1360. https://doi.org/10.1016/j.protcy.2014.10.152

Shih, P.-H. and Chi, K.-H. (2019). "A deep learning application for detecting facade tile degradation", In *Proc. Int'l Conf. Human Syst. Engr. and Design: Future Trends and Applications*, 26-32. https://doi.org/10.1007/978-3-030-27928-8_5

Yasen, M. and Jusoh, S. (2019). "A systematic review on hand gesture recognition techniques, challenges and applications", *PeerJ Comput. Sci*, **5**:e218 http://doi.org/10.7717/peerj-cs.218

Zhang, X., Yang, Z., Chen, T., Chen, D., and Huang, M.-C. (2019), "Cooperative sensing and wearable computing for sequential hand gesture recognition", *IEEE Sensors J*, **19**(14), 5775-5783.