GESTURE RECOGNITION FOR AUTOMATION

SOHAM KORGAONKAR

DHAVAL VIJAPURI

TANMAY KULKARNI
Email ID: soham.sk.sk@gmail.com

Abstract
In this paper we have discussed about Gesture Technology and its implementation using Smart Phones and Smart Watches which are quite easily available, cheap and very easy to use. These devices have sensors built into them, such as accelerometer, which are not used to their full potential. We will use these sensors to recognize gestures and interpret them using microprocessors. These sensors will transfer the information collected by it through Wi-Fi to the microprocessor. This information will be sent to an online server where it will be processed using Artificial Intelligence (A.I). Here the raw digital signals will be converted into required instructions. These instructions will be sent back to the microprocessor and the microprocessor will thus execute the desired task. This project has numerous applications in daily life but most prominent application is for the differently abled people as these people will be benefited a lot with this technology. Other applications may include home automation and navigation.
***Keywords**: gesture recognition, deep learning, accelerometer, automation, microprocessor*

Introduction

Gesture is the physical movement of a body part to express an idea or feeling. These can be recorded by sensors present in phones, smart watches and recognized by Artificial Intelligence (AI). This information can be used for automation purposes.

In today's world each of us has a smart phone and most of us possess a smart watch. These smart phones and watches possess sensors like accelerometer, gyroscope, motion detector etc. With the combined interaction of these sensors, along with Deep Learning AI Technology, we can determine the approximate posture of our body and we can use this data obtained for gesture identification. With the help of this we can automate most of the basic and intermediate tasks in our lives. This technology will be of great use to differently abled people who find it difficult to carry out simple tasks like switching on or switching off lights and fans. With the help of this technology these people will be able to carry out their day to day activities easily and help them to lead a more comfortable life. In India most of the people find this technology inconvenient as the current devices and equipments available for this purpose are either costly or highly sophisticated, making them difficult to use. Due to this inconvenience use of gesture technology is not very widespread. The devices that will be used in this project are very cheap and are used by majority of the people. Hence this project will make gesture technology cheaper and will help promote the widespread use of the same.

Literary survey

[1] Today everyone has smart phones which have good quality camera. Using these phones, microcontrollers and Bluetooth, an efficient gesture control can be implemented. To identify the gestures, the digital signals of the corresponding gestures are stored in the RAM of the mobile. Then

when we do some gestures in front of the mobile camera, the image of our gesture is captured and is processed by Principal Component Analysis (PCA). Thus, the image is converted into a digital signal. Now this digital signal is matched with the previously stored signals in the RAM. If it matches, the binary signal is received by the microcontroller via Bluetooth. Then the microcontroller creates a binary combination to pass it to the decoder. The decoder identifies the code and performs the necessary operation. For this project, the camera of your mobile should at high contrast and your mobile camera should have a flash in it to improve the accuracy of recognising the gestures. [2] Automating the home appliances is known as home automation. The home appliances may involve lights, fans, heaters, AC, etc. Our gestures can be mapped for controlling the home appliances. An interface can be used to connect them and they can be controlled creating a Personal Area Network (PAN). Then these devices can be connected to internet with the help of a router (IOT). An accelerometer can be used to capture the gestures. The voltage output generated by the accelerometer is received by the microcontroller. This microcontroller processes the voltage and sends it to encoder. The work of the encoder is to convert the data into the required form for transmission. Here the data is transferred to the receiver with the help of RF module. At the receiving side, there is a decoder which will decode the received signals. These processed signals will then be sent back to the microcontroller which will then actuate the relays and will switch the appliances. By this way the gesture controlling will be achieved. [3] User Centered Gesture is a gesture technology that involves the user in its devlopment. This makes the technology natural and easy to use. Due to increase in use of computing devices, a controller free interface is needed and use of gesture fulfils it.This technology will have lots of uses in the medical field.A surgeon can use this technology during an operation as the mouse is inaccessible at this time.Noldus Observer X9 can be used for analysis of these gestures.Two camera (stereo) setup is more accurate and is thus suitable for this project.Software for this technology uses small modules which are specialized for detection of a particular gesture.The software can also learn and improve upon the recognition of gestures.Study shows that use of gesture is more tiring than using mouse. [4] Gestures of our finger tips can also be used for the purpose of automation. Initially, we must extract the real time images of the fingertip. Extracting only the image of the fingertip and subtracting the background is often not very accurate when the background contains dynamic lighting. So, an infrared camera is used. This camera is adjusted to measure the temperatures which lie nearby the human body temperature range. Because of this the pixel value of the skin colour gets prioritised over the other colours, making the tracking of the fingertip more accurate. Once the fingertip is identified, then we obtain the trajectory of the fingertip by evaluating the successive image frames. We can do the same thing for multiple fingers as well. From the obtained trajectories, now we can interpret the gestures performed by fingertips. Now these gestures can be transferred to microcontrollers which will then process the gestures into binarized format and then will execute a desired task. [5] Use of hand gestures is a very powerful method in gesture technology. Many devices can be used for this purpose. One such device is the MYO armband. It is a device used for recognising gestures and we have to wear it on our forearm. This is used for using technology wirelessly with the help of our hand movements. It uses EMG sensors. These sensors are responsible to measure electrical activity from our muscles. Then this information is transferred to a compatible device using wireless Bluetooth technology. It uses a combination of gyroscope, accelerometer and magnetometer to identify gestures. There are a certain loopholes in MYO such as, when we place our one hand above the other, it is not able to identify the hand and the fingers accurately. Also, the device is quite costly. MYO armband has its applications

in gaming, presentations and visual entertainment. [6] This paper focuses on the Time Of flight technology (TOF). There are specialised TOF cameras available for this purpose. It uses optic reference signals which are useful to illuminate the things present in front of it. Now using smart pixels, the phase shift, with respect to the reference signal, is estimated. Along with the phase delay, amplitude of the reflected wave is also measured. Thus, with the help of this information we obtain a gray scaled image of the object. Thresholding of the distance and amplitude information is done to improve the accuracy. The reflectivity of this material is encoded in gray values. To decode these values, a decoder is required. [7] The technology suggested in this paper uses a pendant with a camera ringed with IR LEDs and IR pass filter over the lens, allowing gestures to be recorded even in the dark. It has a similar design to Toshiba "Motion Processor" project, which is used for interaction with desktops and portable computers. Since sensing and computing units are on the body, the same pendant can be used as a control unit in multiple locations with ease. Using a variety of complementing elements to increase the number of information sources the number and complexity of gestures can be reduced. It has much higher accuracy than voice recognition systems, and can be easily used in noisy environments as it is unaffected by the ambient noise. A camera with a 160° field of view was used to take the video of the gesture performed by the user. It is simple to use and can be used by elderly people or those with loss of vision, motor skills and mobility. The setup is currently power inefficient due to off-the-shelf components which consume a lot of power. It is also difficult to create a battery with enough capacity for a full day's use and small enough to be lightweight and non-obtrusive. [8] Uses wireless signals for the recognition of gestures, requiring no instrumentation or cameras. Since the signals can pass through walls and do not require line-of-sight communication, a limited number of signal sources can be used for automating a large area. The gestures are recognized using minute doppler shifts and multi-path distortions produced due to interaction of the signals and human motions. Since the doppler shifts produced due to interaction in normal Wi-Fi signals are negligible as compared to the frequency of the signals, the received signal is transformed into a narrowband pulse, making it easier to detect the small doppler shifts. WiSee uses MIMO capability to focus on gestures from a particular user. Using a trigger approach, similar to the ones used by multiplayer games using Xbox Kinect, the user controls the interface by performing a particular gesture pattern. Limitation of WiSee is that the accuracy reduces the number of users increase for a fixed number of transmitters and receiving antennas. [9] Uses a range camera to obtain depth data of an object. 3D range camera captures data unambiguously and at a high frame rate. This eases the task of segmentation and tracking in a 3D space. The 3D camera uses a Time-Of-Flight approach to obtain the distance of a point. This is achieved by using modulated infrared light. Both the depth and intensity data received from the sensor has a lot of noise, making pre-processing essential. This is done by removing those pixels whose amplitude after reflection is below a pre-determined threshold value. The depth is then recalculated at these points using a linear interpolation method. However, points near the edge of the hand end up getting merged into the background as the resolution of the sensor used is limited. Lighting conditions also affect the depth data, but this is rectified by frequent calculation of the average intensity value of the obtained images. The depth value of the object cannot be measured accurately at high speeds, so the data obtained from the centroid is used as a rough estimate of the depth. The hand can be segmented using a histogram method, which compares the collected data and compares to find the section with the least distance and having a sufficient number of points. [10] Uses a 5-sensor data glove to efficiently recognize hand gestures. This approach uses customized gloves which is a cheaper

alternative than vision-based methods which require sophisticated equipment and extensive pre-processing. The gestures used for testing were static only. PNN is the optimal solution for the core component. It has high recognition rates, can be used to accommodate various hand anatomies and can store around 16 hand gestures. However, it uses a lot of memory as the execution is slow. Thus, a combination of clustering algorithms are used along with a four-tier architecture. Every combination of open/closed finger is given an index starting from 0(fist) to 15(open hand). The gestures are recognized on the basis of a boundary value. If the reading obtained for the finger is above this value, it is considered to be closed, if it's above the boundary value, it is considered open. At the boundary value the reading is considered as undefined. This method sometimes misinterprets the gesture due to the anatomical differences between people and cross coupling of optical sensors. Since the simple gestures lacked symbolic meaning, complex static gestures were introduced. This data was recorded using a small Python routine. It saved a record containing six fields, the first five for the readings obtained from the sensor and the last one for the gesture.

Findings

After our study, we found out that not only the sensors in our phones and watches, but also many other devices can be used for this purpose. The devices like MOY armbands, TOF technology can be useful in this project. But these devices are quite expensive as well. The A.I Technology involved in this process can be difficult to develop, but use of online services like Google Development Platform, which allows us to integrate Google's APIs can make this task easier. Currently a lot of research work is going on in this area and there is a lot of scope for future research. But the only thing that makes our project different from others is that, it will be cheaper than most of the current projects. Also, the use of simple devices such as mobile phones and smartwatches is currently unexplored. Due to the introduction of easily accessible cloud services and internet access, it is now possible to connect this project to the internet easily.

Conclusion

Initially, our idea was to control and recognise gestures using only the sensors present in our mobiles and smart watches and the use this gesture technology in the field of automation. After studying about this topic in detail, we realised that the mobile cameras can also be used for capturing the gestures and this would be of great help. A merit of this project is that it would very cheap as it uses smart phones and smart watches which are available at very cheap rates these days. A microprocessor chip required is also a very cheap device. Also, gesture technology is gaining a lot of importance these days and it has its applications in almost all sectors. The demerits of the project include, the range for operation may be limited. The biggest difficulty will be to make it accessible to all. As the field of gesture technology is new, there are a lot of things to investigate. Also, there is a lot of scope for further research in this area. The field of gesture technology would definitely contribute in building energy efficient automated devices.

References

Dey, S., Das, A., Mishra, A., & Dutta, D. (2015). Gesture controlled home automation. International Journal, 162.

Li, Z., & Jarvis, R. (2009, December). Real time hand gesture recognition using a range camera. In *Australasian Conference on Robotics and Automation* (pp. 21-27).

Luzanin, O., & Plancak, M. (2014). Hand gesture recognition using low-budget data glove and cluster-trained probabilistic neural network. *Assembly Automation*, *34*(1), 94-105.

Mulling, T., &Sathiyanarayanan, M. (2015, July). Characteristics of hand gesture navigation: a case study using a wearable device (MYO). In *Proceedings of the 2015 British HCI Conference* (pp. 283-284). ACM.

Oka, K., Sato, Y., & Koike, H. (2002). Real-time fingertip tracking and gesture recognition. *IEEE Computer graphics and Applications*, *22*(6), 64-71.

Penne, J., Soutschek, S., Fedorowicz, L., &Hornegger, J. (2008, September). Robust Real-Time 3D Time-of-Flight Based Gesture Navigation. In *Mensch & Computer Workshopband* (pp. 79-81).

Pu, Q., Gupta, S., Gollakota, S., & Patel, S. (2013, September). Whole-home gesture recognition using wireless signals. In *Proceedings of the 19th annual international conference on Mobile computing & networking* (pp. 27-38). ACM.

Starner, T., Auxier, J., Ashbrook, D., & Gandy, M. (2000, October). The gesture pendant: A self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. In *Wearable computers, the fourth international symposium on* (pp. 87-94). IEEE.

Tambe, P., Nigam, A., Raj, M., & Chaturvedi, R. GESTURE CONTROLLED HOME AUTOMATION.

Van Beurdena, M. H., IJsselsteijna, W. A., & Hopfb, K. (2011, May). User centered design of gesture-based interaction technology. In *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2011* (pp. 1-4). IEEE.