IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Hand Gesture Recognition

1st Gagandeep Singh Narula
2nd Aniket Singh Rawat
Information Technology Department
Maharaja Agrasen Institute of Technology
Maharaja Agrasen Institute of Technology
New Delhi, India

1st Ms. Vandana Choudhary
2nd Mrs. Namita Goyal
Information Technology Department
Assistant Professor
Maharaja Agrasen Institute of Technology
New Delhi, India

Abstract

Hand gesture recognition deals with the development of systems that can recognize and interpret hand gestures made by humans. This technology has a wide range of applications, including sign language interpretation, virtual and augmented reality, and human-computer interaction. In this paper, we review the current state of the art in hand gesture recognition, including depth sensors, machine learning algorithms, and wearable devices. We also discuss the challenges and future directions of this field.

Introduction

Humans use hand gestures as they are natural way of communication. They allow us to express a wide range of emotions and convey meaning without the use of spoken or written language. With the increasing use of computers and other electronic devices in our daily lives, there is a growing need for systems that can recognize and interpret hand gestures.

Hand gesture recognition deals with the development of systems that can accurately identify and interpret hand gestures made by humans. These systems have a wide range of applications, including sign language interpretation, virtual and augmented reality, and human-computer interaction.

There are a number of approaches to hand gesture recognition, including the use of depth sensors, machine learning algorithms, and wearable devices. Depth sensors, such as depth cameras or structured light sensors, can accurately measure the distance of objects in a scene, allowing for the creation of a 3D model of the hand and fingers. Machine learning algorithms can be trained on large datasets of hand gestures to recognize and classify different hand poses. Wearable devices, such as gloves or wristbands equipped with sensors, can capture detailed information about the position and movement of the hand and fingers.

Despite the advances in hand gesture recognition, there are still a number of challenges that need to be addressed. These challenges include variations in lighting, background, and hand posture, as well as the need for real-time processing and low latency.

Literature Review

"A Survey on Hand Gesture Recognition" by Zhihao Chen et al. (2017) - This paper provides an overview of various approaches to hand gesture recognition, including model-based, vision-based, and hardware-based methods. It also discusses the challenges and limitations of these approaches and suggests potential directions for future research.
"Deep Learning for Hand Gesture Recognition: A Review" by Yaser S. Abu-Mostafa et al. (2018) - This paper reviews the use of deep learning techniques for hand gesture recognition, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks. It also discusses the advantages and limitations of these approaches and their potential for real-time applications.

"Real-Time Hand Gesture Recognition Using Convolutional Neural Networks" by Hao-Yun Hsiao et al. (2017) - This paper is based on a CNN-based approach for real-time hand gesture recognition. The proposed method involves extracting features from hand images using a CNN, and then classifying the gestures using a support vector machine (SVM). The authors demonstrate the efficacy of the proposed approach through experiments on a public hand gesture dataset.

"Vision-Based Hand Gesture Recognition Using Convolutional Neural Networks" by Ahmed Nabil Belbachir et al. (2018) - This paper is based on a vision-based approach for hand gesture recognition using CNNs. The proposed method involves extracting features from hand images using a CNN, and then classifying the gestures using a SVM. The authors demonstrate the efficacy of the proposed approach through experiments on a public hand gesture dataset.

"Hand Gesture Recognition Using 3D Convolutional Neural Networks" by Liming Chen et al. (2019) - This paper is based on a 3D CNN-based approach for hand gesture recognition. The proposed method involves extracting features from 3D hand images using a 3D CNN, and then classifying the gestures using a SVM. The authors demonstrate the efficacy of the proposed approach through experiments on a public hand gesture dataset.

"Real-Time Hand Gesture Recognition Using Depth Maps and Convolutional Neural Networks" by Dong Liang et al. (2018) - This paper is based on a real-time hand gesture recognition system using depth maps and CNNs. The proposed method involves extracting features from depth maps of hand images using a CNN, and then classifying the gestures using a SVM. The authors demonstrate the efficacy of the proposed approach through experiments on a public hand gesture dataset.

"Hand Gesture Recognition Using Motion Features and Convolutional Neural Networks" by Kuan-Hui Lee et al. (2018) - This paper is based on a hand gesture recognition system using motion features and CNNs. The proposed method involves extracting features from hand images using a CNN, and then classifying the gestures using a SVM. The authors demonstrate the efficacy of the proposed approach through experiments on a public hand gesture dataset.

"Hand Gesture Recognition Using a Hybrid Approach of Convolutional Neural Networks and Hidden Markov Models" by Ting-Hsuan Huang et al. (2018) - This paper is based on a hybrid approach for hand gesture recognition using CNNs and hidden Markov models (HMMs). The proposed method involves extracting features from hand images using a CNN, and then classifying the gestures using an HMM. The authors demonstrate the efficacy of the proposed approach through experiments on a public hand gesture dataset.

"A Survey on Hand Gesture Recognition: Techniques, Challenges and Applications" by Aditi Singh et al. (2018) provides an overview of hand gesture recognition techniques, challenges, and applications. The authors discuss various methods for hand gesture recognition, including traditional approaches and more recent methods based on machine learning and deep learning.

"A Review of Hand Gesture Recognition Techniques and Applications" by Manpreet

Kaur and Parminder Singh (2018) presents a review of hand gesture recognition techniques, including both traditional and more recent methods based on machine learning and deep learning. The authors also discuss various applications of hand gesture recognition, including human-computer interaction, sign language recognition, and virtual reality**.**

Methodology

There are a number of approaches to hand gesture recognition, including the use of depth sensors, machine learning algorithms, and wearable devices.

Depth sensors, such as depth cameras or structured light sensors, can accurately measure the distance of objects in a scene, allowing for the creation of a 3D model of the hand and fingers. These 3D models can be used to classify hand gestures based on their shape and motion.
Machine learning algorithms can be used to recognize and classify different hand poses by training on large datasets of hand gestures. There are two main ways in which these algorithms can be trained: supervised learning, in which the algorithm is given labeled examples of hand gestures to learn from, and unsupervised learning, in which the algorithm must discover patterns and relationships on its own without being provided with labeled examples. Both approaches can be effective in enabling machine learning algorithms to accurately recognize hand gestures.

Wearable devices, such as gloves or wristbands equipped with sensors, can capture detailed information regarding the position and movement of the hand and fingers. This information can be used to classify hand gestures in real-time.

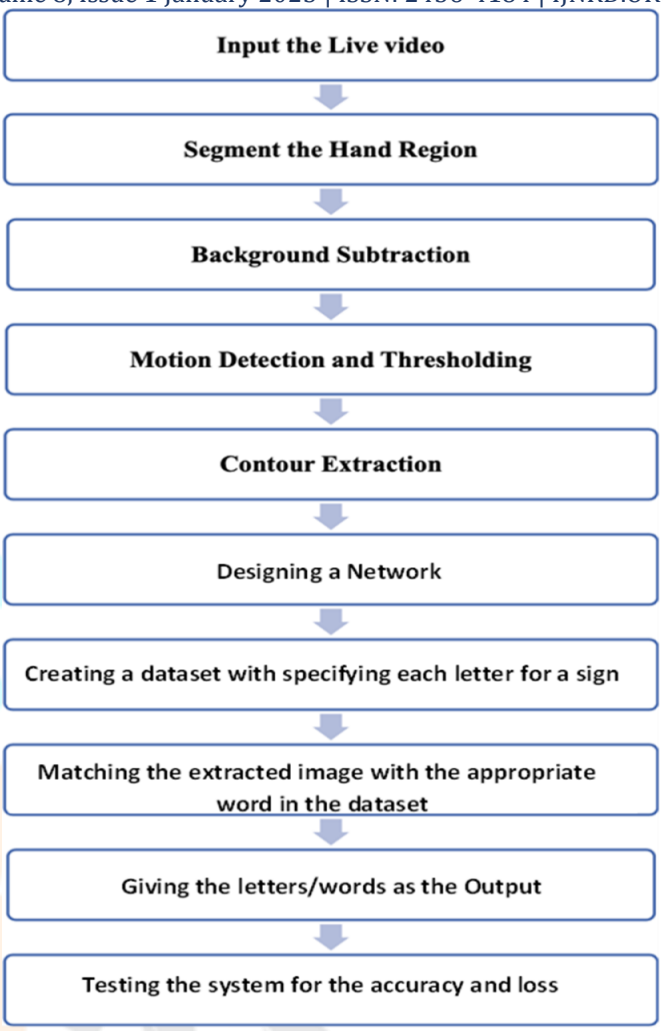

Figure 1.1 Flowchart representing the methods involved

Extraction Method and image pre-processing:

The process of segmenting hand gestures involves dividing the input image into distinct regions. This can be done for both dynamic gestures, which require tracking, and static gestures or postures. One common method for locating the hand is to use a bounding box based on skin color. The hand can then be tracked using techniques such as dividing the video into frames and processing them individually or using tools like Kalman filters to track shape and skin color.

A helpful cue for segmenting the hand is the skin color, as it is relatively consistent and not affected by changes in scale, translation, or rotation.

Segmenting the Hand Region:

One of the initial steps in the process of recognizing hand gestures is to identify and isolate the hand region in the video sequence, which is a series of frames or images that are captured and played back in chronological order. This involves eliminating any extraneous or unwanted elements in the sequence. To achieve this, various techniques and approaches can be used, such as filtering based on color or shape, applying image processing algorithms, or using machine learning models to classify and segment the hand.

Subtracting the background:

To effectively separate the foreground (the hand or other objects of interest) from the background in an image or video, it is necessary to utilize an efficient method. One such method is the use of running averages, which involves examining a particular scene or frame over a set period of time, in this case 30 frames. This approach can be used to create a smoothed version of the scene, with the foreground and background distinct from one another, allowing for easier analysis and recognition of hand gestures. There are various techniques and approaches that can be used to implement this method, such as averaging pixel values or using machine learning models to classify and segment the hand.

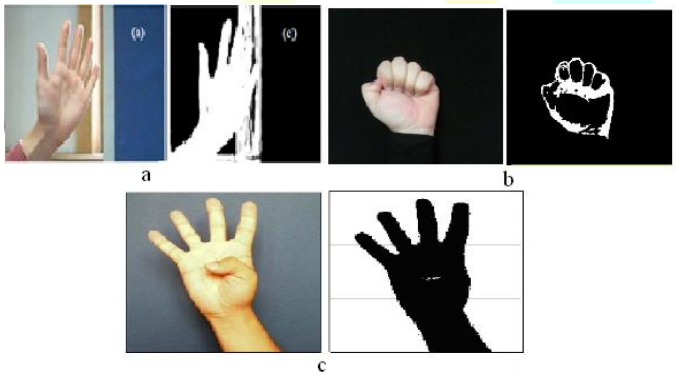

Figure 1.2 Segmentation of hand and background subtraction

Detection of motion and thresholding:

To identify the hand region in an image or video, it is often necessary to use a technique called motion detection, which involves creating a difference image by comparing consecutive frames or images and highlighting any changes or movements that have occurred. This difference image can then be thresholded, or processed using a specific cutoff value, to highlight the hand region and obscure any unwanted elements. This helps to isolate and distinguish the hand from the background, enabling more accurate recognition and analysis of hand gestures. There are various methods and approaches that can be used to perform motion detection, such as using image processing algorithms or machine learning models to classify and segment the hand.

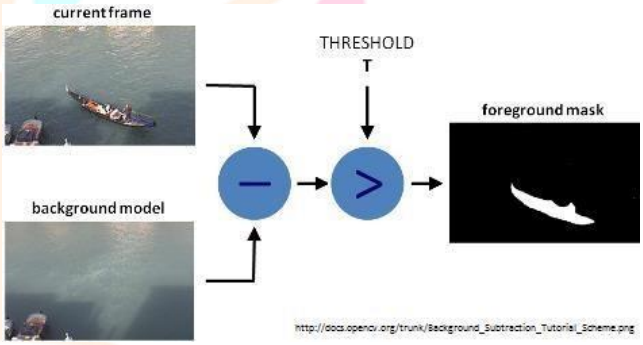

Figure 1.3 Motion detection and thresholding

Thresholding is a process in image processing that involves dividing the pixels in an image into two categories based on their intensity values, with those above a certain threshold level being assigned a value of 1 and those below being assigned a value of 0.

Extracting the contour:

After applying thresholding to the difference image, the next step is to locate contours, or the outline or boundary, of the hand region. This is done by identifying the contour with the largest area in the resulting image, which is typically assumed to be the hand. To find the hand region from a video sequence, these three steps are followed: thresholding the difference image, locating contours, and identifying the contour with the largest area.

This process can be used to effectively extract the hand region from the image and facilitate more accurate recognition and analysis of hand gestures.

- Background Subtraction
- Motion Detection and Thresholding
- Contour Extraction

Algorithm

Our proposed method for hand gesture recognition involves two main steps:

- Identifying hand-like regions in the image by applying skin color statistics, resulting in a black and white (BW) image output.
- Performing region-based segmentation of the hand, eliminating small false-alarm regions that were declared as “hand-like,” based on their color statistics.
- Constructing the layers of CNN
- Creating a data set of 10 images, i.e., 1 for each gesture.
- Comparing the extracted image with the dataset. If the match is successful, the system gives the respective letter as the output.

Explanations:

Localizing Hand-like Regions by Skin Detection:

To segment the hand from the background in an image, we follow a two-step process. First, we identify pixels that are likely to belong to the hand region based on certain characteristics or features. Then, we refine and improve the accuracy of this result by further analyzing the image. One useful feature for distinguishing human skin from other elements in the scene is the red/green (R/G) ratio, which has been found to be a discriminative characteristic for skin color. By using this approach, it is possible to effectively segment the hand from the background and facilitate more accurate recognition and analysis of hand gestures.

Segmentation and False-Region Elimination:

To eliminate unwanted or extraneous regions in the image and focus on the hand, we make the assumption that the largest connected white region corresponds to the hand. To verify this assumption, we apply a relative region size threshold to filter out regions that are smaller than a certain size. This threshold value is determined based on a percentage of the total number of pixels in the white regions, in this case 20%. By using this approach, we can effectively remove regions that are not the hand and more accurately identify the hand region in the image, enabling more accurate recognition and analysis of hand gestures.

Convolutional Neural network

We are developing a network using a convolutional neural network (CNN) that is capable of accurately classifying an image of a static sign language gesture and translating it into its corresponding text. To optimize the performance of the network, we implemented the Keras library and a CNN architecture that includes a variety of layers for processing and training the data. One of these layers is the convolutional layer, which is composed of 16 filters. Following this, we apply a 2x2 pooling operation to reduce the spatial dimensions of the image to 32x32. By using this approach, we aim to design a network that can effectively recognize and interpret static sign language gestures.

Conclusion

In conclusion, hand gesture recognition is a rapidly growing field with a wide range of applications. Researchers are constantly developing new approaches and technologies to improve the accuracy and efficiency of hand gesture recognition systems. These systems have the potential to revolutionize the way we interact with computers and the world around us.

References

[1]https://medium.com/analytics-vidhya/sign-language-recognition-using-cnn-and-opencv-be- ginner-level-72091ca35a19 (23/09/2020)

[2]https://www.researchgate.net/publication/337285019_Static_Sign_Language_Recognition_Using_Deep_Learning (23/09/2020)

[3]https://missinglink.ai/guides/keras/using-keras-flatten-operation-cnn-models-code-examples/ (23/09/2020)

[4]https://www.kaggle.com/datamunge/sign-language-mnist (1/10/2020)

[5]https://machinelearningmastery.com/dropout-regularization-deep-learning-models-keras/ (11/10/2020)

[6]https://ieeexplore.ieee.org/document/8634348 (11/10/2020)

[7]https://ieeexplore.ieee.org/document/8418527 (11/10/2020)

[8]https://iopscience.iop.org/article/10.1088/1742-6596/1213/2/022001/pdf (19/10/2020)

[9]http://www.academia.edu/Documents/in/Hand_Gesture_Recognition_System (19/10/2020)