Texas A&M University Qatar

ECEN 403

Functional Assignment

Project Title:

Gesture Guide

Instructor:

Dr. Khalid Qaraqe

Mentor:

Dr. Hussein Al-Nuweiri

Team 3:

Maha Al-Dehemi

Fatma Al-Mohannadi

Maryam Al-Safran

Lolwa Al-Kaabi

Due Date: 26/05/2024

Texas A&M University at Qatar, 2024

“An Aggie does not lie, cheat or steal or tolerate those who do.”

Contents

1. Introduction:

The aim of the project design is to help people with physical disabilities to control their environment. In order to put the system into effect efficiently and help people with disabilities to continue their lives with minimal effort using our project and access basic needs, we hope to do these using two versions of the project that will be designed, built and tested.

The first device is a monitoring system connected to a 3D prosthetic hand. It will display any finger movement as a receiver on the prosthetic hand as a similar action thunder.

The second version will be connected to a set of different tasks that the person can define himself according to need, such as calling a specific person or calling the attendant for the patient or opening or closing the room lights or controlling the air conditioner or opening or closing the curtains and other multiple tasks that can be defined according to the need and environment of each person.

2. Upper Level Functional Modeling:

In order to delve deeper into the gesture system, we must first distinguish between the meaning of a pose and a gesture. A pose is a single image that represents a single command, such as a stop sign, while a gesture is a series of poses that indicate a unique meaning when these poses are combined together, such as moving the hand in a certain direction to change the volume of a radio or television. The hand must be modeled in the system in order to be processed correctly.

Gesture recognition system composed of several stages; these stages are varied according to application used, but, however, the unified outline can be settled, Figure 1 fulfils this step.

Fig1: Functional Modeling

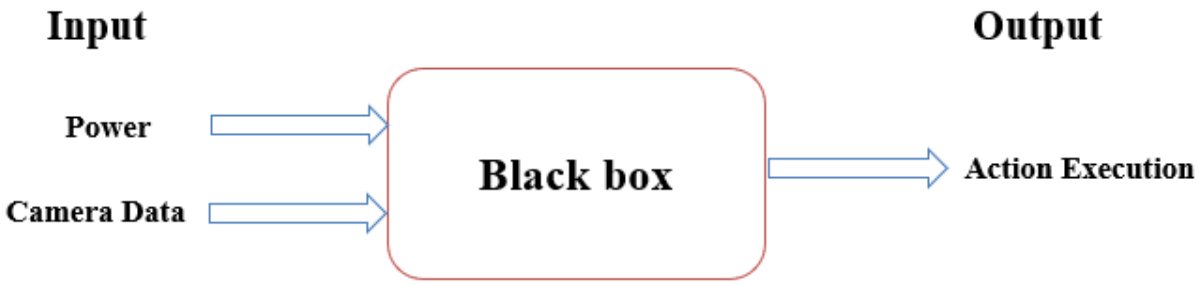

Fig2: Block Diagram

The application requirement has a great influence on the choice of model used. The hand model can be temporal (motion), and special recognizers have been created to track temporal modeling, such as the Hidden Markov Model (HMM) [1], Neural Network (NN), Rule Machine-based and Finite State Machine [2].

The special modeling can be divided into appearance-based model and 3D-based model. The appearance-based model is also referred to as the 2D model or view-based model [1].

The appearance-based model can be formed by templates, representation features, and eigenvectors. The shape representation features can be either geometric or non-geometric features. Geometric features are considered as live features as they can be processed separately such as fingertip locations and palm location. On the contrary, non-geometric features are considered as blind features as they cannot be seen individually and require collective processing.

The 3D model describes the shape of the hand and can be divided into volumetric models and structural models. Volumetric models are complex to implement especially in real-time applications where the speed factor is critical.

Therefore, other geometric models such as cylinders and spheres are considered as alternatives to such a model to approximate the shape of the hand. The other type of model is the structural model which captures the structure of the hand in 3D structure with a reduced set of parameters compared to the previous model. When a volumetric model is cast on an appearance-based model, the result is complex and time-consuming to calculate the parameters and the approach at this time can be called analysis by synthesis.

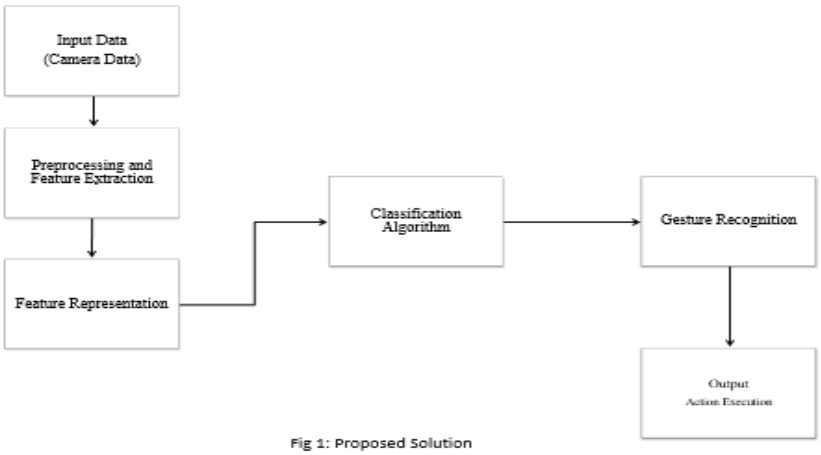

Fig 1: Proposed Solution

Figure 2 shows a pictorial representation of the above classification. This classification is done through different categories based on the interaction of hand movement which can take one of the following two types: direct manipulation and symbolic gestures. In direct manipulation; The current hand position can be interpreted as the next command while the symbolic can be obtained from the object motion. However, 3D modeling solves the self-occlusion problem but is not useful for real-time applications because time is against this modeling.

Moreover, both 3D and 2D models have different parameters that can be used for model estimation, feature vector representation and correct recognition, these parameters can overlap between these two models, joint angles, spatial position of palm, and spatial orientation have been extensively used for 3D model because it can capture the hand information, on the other hand, image can be used as input for non-geometric features, geometric features, and fingertip locations for a 2D vision based system. However, for tracking information; we can use hand motion and fingertip motion information for this purpose. Joint angles can be modeled as a whole arm namely shoulder, elbow, and wrist, or it can refer to the finger angle, however, since the joint property of the hand body has a flexible degree of freedom it is difficult to model it perfectly in the system.

3. Detailed functional modelling:

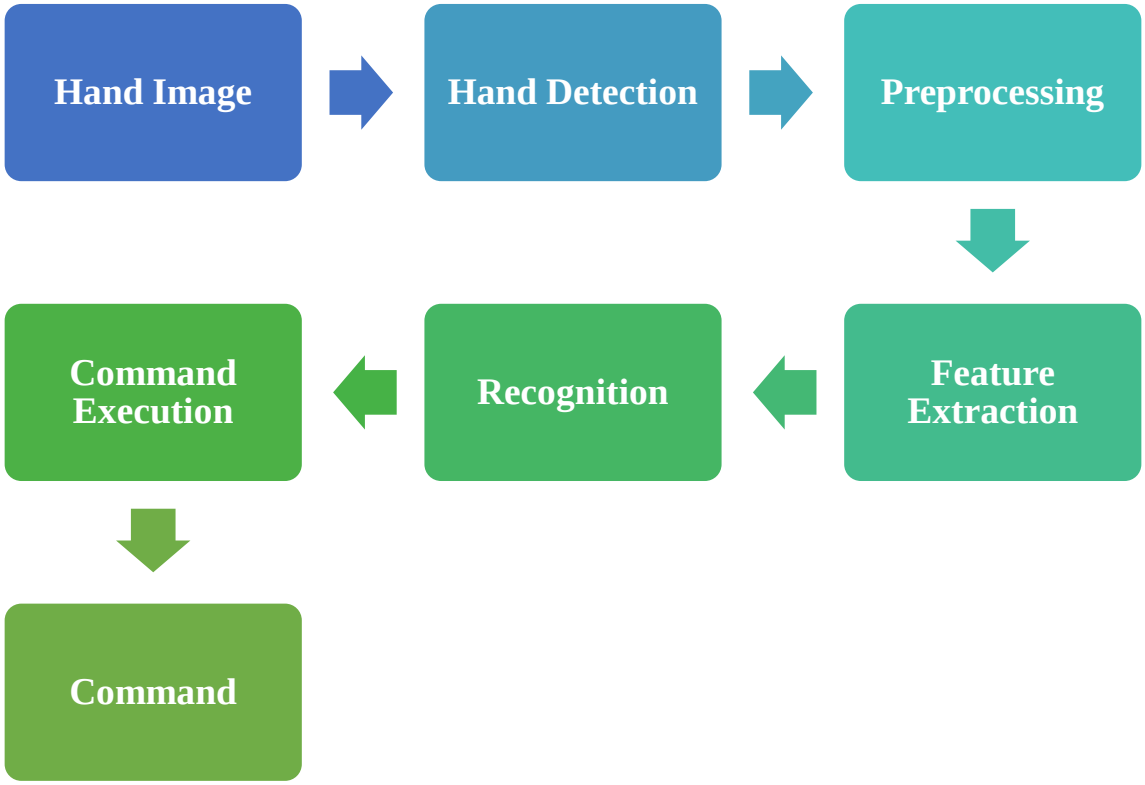

Fig 3: Suggested Process

The suggested Process of hand gesture has the following steps:

1. **Hand Image:**
   From the recorded video stream, extract a frame, that is, a hand image.

2. **Hand Detection:**
   The extracted frame is converted from the color space of RGB to the color space model of YCbCr. Then using skin color-based detection techniques, the hand is detected in the image.

3. **Preprocessing:**
   The machine transformed the picture into black and white after hand recognition (i.e., the skin pixels were identified as white and nonskin pixels as black). Some preprocessing strategies, such as picture filling,

morphological erosion utilizing 15×15 structuring elements, etc., were implemented to enhance image clarity and eliminate noise.

4. **Feature Extraction:**
   The equivalent diameter, field, perimeter, and orientation of detected objects are found in the frame for the function extraction centroid. All the features were used before we had the no conflicting production.

5. **Recognition:**
   The gesture is recognized by counting the number of white items in the picture and its direction. Lastly, an instruction is transmitted to the device's programs, referring to the known motion.

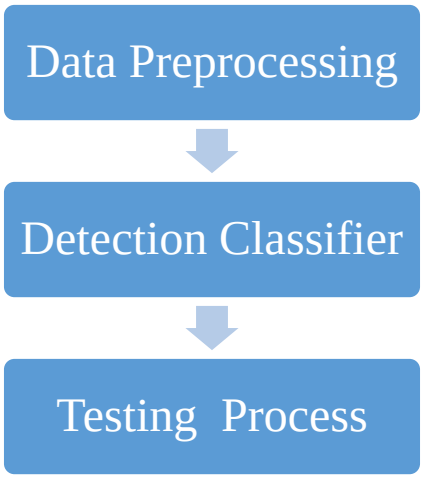

Fig 4: Object recognition

Classifiers based on features are a machine-based learning method in which both positive and negative photos learn a function. It's then used to classify items in other pictures. Objects are drawn from both positive and negative photos in this method, as seen in Fig.3. The positive picture is an image having just the right thing. The gloomy picture is an illustration that doesn't have the necessary individual. In cascades' production, a favorable image is taken, and

negative images are mounted on it. In this method, positive images are resized and distorted at different angles into negative images.

6. **Execute orders:**
   As we mentioned earlier, the orders that each patient wants differ from one person to another, and therefore this stage depends on the orders that the patient wants from the project, as we mentioned earlier.

7. **Command:**
   This is the last stage, which is to implement the required order by calling or executing a specific order.

4. Conclusion:

Hand gesture recognition system has become an important role in building effective human-machine interaction. The implementation using hand gesture recognition is widely used in the technology industry. The project as a single framework based on machine learning plays an effective role in developing this application using hand gesture recognition, with previous results showing 95% accuracy performance. We would like to further expand our system to develop cooperation with other devices and other parts of the human body and experiment with static and dynamic hand gesture recognition systems.

5. References:

1- Ustunug A, Cevikcan, Industry 4.0: Managing The Digital Transformation, Springer Series in Advanced Manufacturing, Switzerland. 2023. DOI: https://doi.org/10.1007/978-3-319-57870-5.

2- Pantic M, Nijholt A, Pentland A, Huanag TS, Human-Centered Intelligent Human-Computer Interaction (HCI2): How Far We From Attaining It?,International Jounal of Autonomous and Adaptive Communications Systems (IJAACS), vol.1 no.2, 2020. pp 168-187.

3- Hamed Al-Saedi A.K, Hassin Al-Asadi A, Survey of Hand Gesture Recognition System. IOP Conferences Series: Journal of Physics: Conferences Series 1294 042003. 2019.

4- C.Chua, H. Guan, Y.Ho, Model-Based 3D Hand Posture Estimation From a Single 2D Image. Image and Vision Computing vol.20, 2022, pp. 191-202.

5- MediaPipe: On-Device, Real Time Hand Tracking, In https://ai.googleblog.com/2019/08/on-devicereal-time-hand-tracking-with.html. 2019. Access 2021.