IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 7.12 ⁜ Vol. 10, Issue 1, January 2023

DOI: 10.17148/IARJSET.2023.10103

TRACKING OF MOBILE PHONES FOR PIRACY DETECTION

Chandana L[1], Danesh Raju V[2], Priyanka K[3], Dheemanth K N[4]

K. S Institute of Technology, Electronics and Communication Engineering, Bangalore [1-4]

Abstract: The illegal copying of movies has been present from many years, to overcome this issue we are here to propose a system that uses IP cameras, Python/Matlab, and a standard general computer to detect the occurrence of piracy and alerting the respective personals about it. This system is expected to be a one-time installment and easily usable by the employee.

Keywords: Piracy, IP Camera, NVR.

I. INTRODUCTION

Gestures are a type of nonverbal or non-speech communication in which preset messages are sent using outwardly visible physical motions in addition to or instead of language. A gesture is any movement of the hand, face, or other bodily part. People may also communicate with robots by using gestures to achieve freehand involvement. Gesture recognition is a branch of computer vision that uses mathematical methods to analyse human gestures. Though gestures can be made with any part of the body, they are most usually made with the hands and the face. The current focus in this research is on the identification of emotions in faces and hand gestures. Users may operate and interact with electronics using straightforward gestures without having to make direct physical contact.

There have been various attempts to use cameras and computer vision algorithms to understand sign language. On the other hand, posture, mobility, intimacy, and human behaviour are identified and recognised using gesture recognition algorithms. Because it confines most input to the keyboard and mouse and enables natural interaction without the need of mechanical devices, it establishes a greater link between computers and humans than the traditional text user interface or his GUI (graphical user interface). With the increasing use of computer interfaces, there are two different sorts of gestures. We're referring to digital motions. Direct adjustments like size and rotation can also be considered online gestures. Offline motions, on the other hand, are normally handled after the interaction has stopped. Offline gestures are those that a user makes after interacting with an object. One example is the start of a menu. Online, direct manipulation gestures are used. These are useful for scaling and rotating physical objects. Depending on the nature of the incoming data, you will interpret the gesture differently. However, the bulk of techniques make use of 3D coordinate-based key pointers. Based on these relative motions, gestures can be recognised with high accuracy, depending on the input quality and algorithmic logic. Body motions must be categorised according to both their common characteristics and the messages they might convey. In sign language, for instance, each gesture represents a word or sentence.

Two approaches to gesture recognition are described in the literature: appearance-based and 3D model-based. The first method calculates a number of crucial metrics, including joint angles or palm position, using 3D data from important body components. However, appearance-based algorithms can quickly decipher images or movies. There are several issues that need to be overcome in order for gesture recognition software to be accurate and useful. There are limitations to both picture noise and the technology used for gesture recognition in images. Photos and videos might not have been taken in the same spot or under the same lighting conditions. The environment or particular user features might make it more difficult to recognise someone. The multiplicity of image-based gesture recognition methods could limit the technology's overall usefulness. For instance, a solution designed for one camera might not work effectively with another. Tracking and identification are made more difficult by the quantity of background noise, particularly when partial and total occlusions happen. The position, resolution, and quality of the camera also have an impact on recognition accuracy. Strong computer vision algorithms are required to capture human motions using visual sensors, such as head movements, facial expressions, and gaze direction.

The detection of hand posture and hand tracking are two of these techniques. In our experiment, movie theatre piracy is found using gesture recognition. By recognising human body movement, we want to employ gesture recognition to locate the perpetrator and identify pirate operations (especially hand movements). For detecting gestures and motions, a variety of library files and methods, including OpenCV and Yolo soon, are available. An IP camera is the type of camera we utilised for this project. Modern cameras have the ability to provide a live feed to the user through the internet. A form

DOI: 10.17148/IARJSET.2023.10103

of digital video camera known as an IP camera, or Internet Protocol camera, receives control data and transmits picture data via an IP network. They are often used for surveillance, unlike analogue closed-circuit television (CCTV) cameras, but all they need is a local area network. The names "IP camera" or "net-cam" are frequently used to denote those that can be seen directly through a network connection, even though the majority of IP cameras are webcams. To control alarms, video recording, and video, certain IP cameras need an NVR. Because the camera may record directly to any adjacent or distant storage media, others can operate decentralizedly.

II. LITERATURE SURVEY

Sánchez-Nielsen et al. [1] Despite the fact that input devices have been around for more than 20 years, many individuals still find using computers to be unpleasant. It is crucial to make attempts to make computers more responsive to nonverbal cues and human speech in its natural form. The PUI paradigm has emerged as a post-WIMP interaction model to satisfy these needs. This study suggests a real-time vision system for use in visual interaction situations through the recognition of hand gestures, utilising general-purpose hardware and affordable sensors, including a simple personal computer and a USB camera, so any user may utilise it in his or her workplace or home.

This technique's base is a rapid segmentation step that can handle a range of hand forms against diverse backdrops and lighting conditions and extracts the moving hand from the whole image. The hand posture is then extracted from the temporal sequence of segmented hands using a recognition algorithm. Using an edge map-based Hausdorff distance approach, a robust shape comparison is the most important phase in the identification process. A visual memory enables the system to accept variations within a gesture and speed up the identification process by storing multiple variables connected to each motion. This page provides experimental evaluations of the 26 hand postures' recognition mechanisms. Experiments show that the system can achieve an average recognition rate of 90% and is suitable for real-time applications. Humans typically use body language to communicate, which may be used in addition to spoken words or even stand alone to convey a whole message. Automatic posture recognition algorithms may be used to improve human-machine interaction [Turk98]. These human-machine interfaces allow a human user to control a variety of devices remotely via hand gestures. [Pen98, Wil95, Ju97, Dam97] Numerous applications, such as contact-less control and home appliances for welfare enhancement, have been proposed. One of the main objectives of this method is to develop a cheap computer vision system that can function on a standard PC equipped with a USB camera. The lighting and scene backdrop complexity levels must be adjustable for the system to function, and they must remain constant throughout execution.

The image processing operators must be kept as time-consuming as feasible to achieve the high processing rate necessary to achieve real-time speed. Additionally, certain operators must be adaptable enough to operate in various lighting and backdrop conditions. The cluster associated with skin colour must be found within the appropriate colour space in order to simulate skin tone. The hue and saturation pair of skin-tone colours are independent of the intensity component, hence the HSI space (Hue, Saturation, and Intensity) was selected [Jon98]. As a result, colours may be defined using just two parameters as opposed to the three needed to describe colours in RGB space. An picture captured with an inexpensive camera can get distorted by arbitrary changes in lighting and intensity. Using a linear smoothing filter, the colours were homogenised and noisy pixels were eliminated. A mean filter offered the best outcomes when compared to the other suggested lineal filtering techniques. [Jai95]. After the original image has been normalised and smoothed, a binary image is produced, where each white pixel substitutes for a skin-tone pixel from the original image. The categorization of skin tones is based on the normalised picture and HSI space colour considerations. Then, a pixel was labelled as a skin tone pixel if its hue and saturation components were within a predetermined range.

However, these ranges do vary a little bit depending on the lighting, the user's skin tone, and the background. These ranges are defined by two rectangles in the HS plane: the R1 rectangle for natural light (0 H 15; 20 S 120) and the R2 rectangle for artificial light (0 H 30; 60 S 160). After the user hand has been segmented, it is advised to detect the hand posture using a model-based method based on the Hausdorff distance that makes use of edge map images and a visual memory. On an AMD K7 700 MHz computer running Windows XP and outfitted with a USB Logitech Quick-cam camera, the hand posture detection and identification method was implemented using the Borland Delphi programming language. Tests of the recognition approach have been conducted using images taken in the real world as the input mechanism. Real-world images have been used as the input method for user interfaces, security systems, autonomous robot vision systems, and applications in virtual worlds while the recognition approach has been assessed. All tests were conducted using 128x96 pictures with a 24b colour depth and a sampling rate of 25 frames per second. A visual memory system, a quick processing method, and a reliable matching process employing the Hausdorff distance technique have all been presented for the problem of hand posture detection and identification. Numerous lighting configurations, backdrops, and human users have been utilised in studies to evaluate the system's performance. The detection rates show how robust the system is against similar postures. Additionally, real-time video applications may be used with a simple

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 7.12 ⋇ Vol. 10, Issue 1, January 2023

DOI: 10.17148/IARJSET.2023.10103

personal computer and a common USB camera thanks to the runtime behaviour. Future work will examine efficient hierarchical N-template matching and explore alternative robust and reliable methods for integrating the system's components into a gesture interface for an anthropomorphic autonomous robot with an active vision system [Her99] and into virtual environment applications.

Ionescu et al. [2] This study explores the use of computer vision to decipher human gestures. Hand gestures would be a natural and straightforward approach to interact with others electronically, direct robots to do specified jobs in hazardous situations, or interface with computers. Hand gestures may be divided into two categories: static gestures and dynamic gestures. This research proposes a novel dynamic hand gesture recognition approach. It is based on an image of a two-dimensional hand skeleton. The hand skeletons from each position are superimposed to form each motion, creating a single picture that serves as the gesture's dynamic characteristic. This signature is compared to the gestures in it using a gesture alphabet, and discrepancies are found by comparing the distance between model parameters, as determined by Baddeley's distance. One can understand how a gesture develops from the first idea to the end performance in a clear motion pattern in both space and time by using hand motions as a model. Kendon distinguishes three motion phases within a single gesture: preparation, stroke, and retraction. The arrangement of the fingers and thumbs in the palm, which determines the hand posture, can distinguish between static and dynamic hand movements (categorized by the initial and final stroke hand configuration and the general stroke motion).

The most organic way is arguably hand gesture analysis, which may be used to develop a human-computer gesture interface. Because of the existing limits in machine vision, it is also the most difficult to apply properly. One or more cameras are used to do the vision-based analysis. The necessary gesture is retrieved after acquiring visual information about a person in a visual environment. Separating a moving hand from potentially complicated surroundings can be difficult when analysing hand motions, tracking hand locations in respect to the environment, and finally identifying hand postures. Some individuals employ markers, marked gloves, or restricted settings like a homogeneous background, a very limited gesture vocabulary, or even a plain posture analysis to assist these processes. A novel technique for identifying dynamic hand movements is put forward in this research. The algorithm is entirely simulated. A single digital video camera is used to capture hand movements, and a computer is used to process the images. The visuals are composed of two-dimensional greyscale images. The hand motions are made in front of a virtually constant background, with a nearly constant distance between the hand and the camera, and in a previously established region of space. Hand segmentation is facilitated by this constrained situation.

The generality of the recognition procedure is unchanged. To maintain the method's universality, no particular gloves or markers are used. Dealing with pixel intensities requires less time than training neural networks or deriving statistics from gestures. The simplest way to compare these characteristics is to determine their exact distance from one another. Before the classification stage, there must also be a brief training phase that comprises a collection of examples for each gesture. The traditional template matching strategy will be used for categorization. All of these factors contribute to effective and accurate hand recognition techniques. As a consequence, the suggested method's overall procedure consists of four basic components. A preparatory step to help the user focus on the gesture A phase in the feature extraction process that combines two techniques: For static recognition, the first method employs histograms of local orientations in the gesture picture (by computing the local gradient on the image). Based on a formula developed by Roth and Freeman. The gesture's feature set will be represented by the orientation histogram. The second, used for dynamic gesture recognition, extracts "the dynamic signature" and is defined using a completely new, unique approach that provides a compressed and efficient representation of the action. For each picture (hand posture) in the gesture image sequence, the dynamic signature is a binary image that shows all of the hand skeletons overlaid on top of one another. Additionally, it shows the gesture's motion information.

The gesture alphabet must be known in order to complete the gesture categorization step of training. There must be a training period. A training set of data is produced for both static and dynamic recognition using a collection of samples for each gesture. Small translations and rotations are invariant when employing different spatial representations of the same gesture. A categorization technique. The features of the training data will be contrasted with the attributes of the unknown gestures. Based on the best match, the unidentified motion will be recognised. The study's author suggests a technique for identifying gestures that makes use of both static and distinguishing dynamic signals. The static signature uses local orientation histograms to categorise hand motions and is based on the approach put out by Freeman and Roth. Since the computation of the orientation histogram is straightforward and can be completed in real time by a computer, the technique is quick and mostly unaffected by changes in light. It would not be appropriate to draw inferences about missing data using this method. To produce the mean orientation histograms, a training set of at least three samples for each gesture in the gesture alphabet is needed. The most current and prior gesture orientation histograms are compared. The training set is compared to the current gesture orientation histogram. A collection of photos is used to symbolise

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 7.12 ⋇ Vol. 10, Issue 1, January 2023

DOI: 10.17148/IARJSET.2023.10103

each gesture. The superposition of all hand region skeletons for each picture in the series acts as the categorization dynamic signature. It represents the gesture's action as well as its geographical placement. For the categorization of gestures, just minimum training is required. There are a number of example signatures calculated for each gesture in the gesture alphabet. On a ten-gesture alphabet, the algorithm performed admirably.

Fang et al. [3] This study offers a technique for real-time hand gesture recognition. To locate the location of the hand, "Skin colour detection" and "Motion detection" can be included. Oriented gradient histograms are used to define hand pictures (HOG). The distance between the project features and the centre of each gesture is calculated using the extracted HOG features in low dimensional subspace, followed by the closest neighbour approach to obtain the recognition results. According to the experimental results, the suggested method had a real-time detection rate of up to 91%. To locate hand regions, adaptive skin colour detection and motion detection are applied.

We have a current overview of skin colour modelling and detection techniques from Kakumanu et al. Another useful concept that is often applied in computer vision to recognise and eliminate regions of interest is motion detection. Using background subtraction, motion may be found in a picture. An identifiable local distribution of edges or gradients of intensity can accurately characterise a hand gesture. The spatial distribution of the local intensity gradient is represented by the HOG feature during HOG feature extraction. A image is first broken into overlapping blocks, which are further divided into smaller, non-overlapping spatial parts known as cells, in order to compute the HOG feature. A local one-dimensional histogram of gradient axes is created by combining all of the pixels in each cell. The HOG feature vector representation is made by concatenating the histogram elements from the cells. PCA-LDA: The HOG feature space might be quite dimensional. A reduction in the computations will save several hundred to several thousand dollars.

The HOG is projected using PCA-LDA. reducing a subspace of 1296 dimensions to 9 dimensions. The linear dimensionality reduction technique used by PCA algorithms. LDA is a method of lowering dimensions where the greatest scatter is obtained for all projected samples. The dispersion is attempted to be shaped to improve classification accuracy. LDA chooses projection matrix W in this manner. The ratio of between-class scatter to within-class scatter determines how to optimise the scatter factor. The original features are initially projected with PCA to an m-dimensional intermediate space, and subsequently with LDA to a final subspace Wss with dimension C-1 (where C is the number of classes). Because there are ten gestures, there are a total of nine subspace dimensions in this work. After projecting all training samples into Wss and computing the mean Wss of each class Xi in this subspace, a classifier may be built using the closest neighbour technique. Assume that in order to be identified, an input sample has to have an xss representation in the subspace. This paper's author presents a method for real-time hand gesture identification. To generate hand position hypotheses, motion detection and adaptive skin colour detection were coupled. The hand areas were described using histograms based on oriented gradients (HOG). Using PCA-LDA, extracted HOG features were projected onto a low-dimensional subspace. A closest neighbour classifier was employed to recognise gestures in the subspace. The testing findings demonstrated the real-time processing capability and up to 91% detection rate of the suggested technique.

Rokade et al. [4] Sign language is the most expressive and natural form of communication for hearing-impaired persons. Its most attractive use is in the creation of more beneficial and friendly human-machine interactions. Gestures are a quick and effective method of communication. A mute person can interact with non-mute persons using a hand gesture recognition system without the use of an interpreter. The author of this research suggests a revolutionary image-based method for hand gesture recognition. Color is important in the process of producing input pictures. Various colour space approaches are used for segmentation. It varies from one to individual and is sensitive to changes in illumination. In the recommended technique, the RGB picture is initially captured by the camera. The photos in this database are large and have a resolution of 768 x 576. The calculation time will be quite long if this size of picture is used for processing. The image is reduced in size to 256 X 192 so that it may be processed further.

The following steps are part of the segmentation process. matching the histogram The colour of the hand is employed in this suggested way to partition the hand. Color cannot be altered by hand rotation or geometrical modifications. Figure 1 depicts the segmentation algorithm. Separates the luminance component in processes I and Q since it results in a number of issues. The input image is altered to produce a YIQ image. Take a look at the result of the grayscale to binary picture conversion. Take the value as the output image's binary conversion cutoff point. A different threshold value will apply based on the illumination. It should be observed that the segmentation results are outstanding, or noiseless, even without extra processing. Cleaning and binary image thinning should be applied to the binary image. Figure 6 displays a binary picture together with a thin binary image. Due to picture thinning, the final image can have some segments that don't match the fingers. Therefore, before detecting the raised fingers, the thin picture has to be cleaned. Track down the joint and end points. The terminal pixel of a thin segment is an end point with just one neighbour and eight connections. A joint point is a thin segment point with more than two 8-connectivity neighbours. Characteristics of gesture recognition:

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 7.12 ⋇ Vol. 10, Issue 1, January 2023

DOI: 10.17148/IARJSET.2023.10103

This approach uses two features to identify gestures. Two characteristics should be calculated for each endpoint. Angle formed by the centre point, the vertical extended line, and the lines connecting each end point to the centre point. The horizontal distance is the separation between an end point and an extended vertical line. The horizontal distance was impacted by the hand's length or the distance the camera was from it. The range of angles and horizontal distances varies for each finger. The tips of the little, ring, middle, and index fingers are located more than 20% higher than the upper line and the first row of the picture. The thumb finger's end point is located above the j-lower and to the right of the centre point. The identification of motions and raised fingers is done using these assumptions. Due to its underlying social relevance and intrinsic complexity, engaging with a deaf person is a highly difficult subject. This method works nicely with ASL static letters. Without the aid of a single individual, the author described a method for identifying hand movements against complex and irregular backdrops. The results of picture segmentation are superior even in low-light conditions. The advantage of the presented method is that post-processing, unlike a hole-filling algorithm, is not required.

Garg et al. [5] Outdated user interface methods like the keyboard, mouse, and pen are no longer relevant thanks to the emergence of ubiquitous computing. These device restrictions have an impact on the command set's usage. From text-based interfaces to 2D graphical interfaces, multimedia-supported interfaces, and fully complete multi-participant Virtual Environment (VE) systems, Natural Human Computer Interaction has advanced. The direct use of hands as an input device is a desirable approach for doing this. Think of a future 3D application where you can rotate and manipulate items by merely moving and turning your hand - without requiring any input devices. This essay discusses how to recognise hand movements visually. The current approaches are divided into 3D model-based and appearance-based approaches, with advantages and disadvantages highlighted and outstanding issues listed. Using a three-dimensional hand model as a basis for an approach Approaches compare input photographs to the possible 2D appearance projected by the 3-D hand model in an effort to predict hand parameters. This is accomplished by comparing the input photos to the highly flexible 3D kinematic hand model (DOFs). This method works well for creating realistic interactions in virtual environments. Rehg and Kanade proposed one of the first model-based solutions to the bare hand tracking problem.

The model-based hand tracking system Digit Eyes, which can at up to 10 Hz recover the state of a 27 DOF hand model from typical grey scale photos, is described in this article. The hand tracking problem is presented in terms of finding the model's fundamental parameters given an image frame (such as an edge map). Due to the trigonometric functions used to simulate joint motions and perspective picture projection, the inverse mapping is nonlinear. It is important to note that when the settings are changed, the final image changes gradually. In appearance-based techniques, the visual characteristics of the hand are modelled using image features, and these parameters are then compared to the image features that were recovered from the video input. Due to the use of less complex 2 D image characteristics, appearance-based techniques have the advantages of real-time performance. Recent years have seen a large number of appearance-based technique research efforts. A simple and well-liked strategy is to look for skin-colored portions of the photograph. Even though this is frequently employed, there are certain drawbacks.

The detection of skin colour is extremely light-sensitive to begin with. Despite the fact that there are excellent and practical methods for detecting skin colour under known and regulated lighting, it is difficult to learn a flexible skin model and change it over time. Second, if there isn't another item in the picture with a texture like skin, this is certainly functioning. Lars and Lindberg identified hand motions using scale-space colour characteristics. Because the existing algorithms are more straightforward than those used for mammalian vision, vision-based hand gesture detection is still a major topic of research. The fundamental issue with most techniques is that, while they may be true in a controlled lab context, they do not hold true in random circumstances. This is because most approaches are based on a variety of underlying presumptions. Two common assumptions are immobile backdrops with high contrast and general lighting conditions. Additionally, the literature relies on the unique data sets of each author to offer recognition findings, making it challenging to compare approaches and casting doubt on their generalizability.

The bulk of the approaches have few features as well. The use of AI to train classifiers is a recent advancement in the identification of hand gestures. It takes a lot of effort to select the features that best represent the item being identified, and training often necessitates a lot of data. Another unsolved issue is identifying the temporal beginning and ending points of meaningful gestures from continuous hand motion (s). This is referred to as "temporal gesture segmentation" or "gesture spotting." Finding ways to reduce training timeframes and develop real-time gesture recognition systems that are affordable, resilient to ambient and lighting conditions, and require no extra hardware is a huge yet interesting research problem. Processing rates have substantially grown in the contemporary digital era, and computers have advanced to the point where they can assist people in difficult jobs. However, some jobs appear to be significantly hampered by input technologies, which underutilize resources and limit the expressiveness of programme use. This is where the ability to recognise hand gestures is useful. Computer vision approaches for hand gesture interfaces must beat present performance in terms of robustness and speed in order to be interactive and usable. It has been examined how to recognise hand

DOI: 10.17148/IARJSET.2023.10103

movements visually. Given how new the science of vision-based gesture recognition is, incredible strides have been achieved toward the eventual objective of enabling human-machine interaction on their terms.

Suarez et al. [6] This study reviews the research on the use of depth in hand tracking and gesture detection. The study looks at 37 papers that talk about depth-based gesture recognition systems in terms of the methods developed and put to use for hand localization and gesture classification, the applications that gesture recognition has been tested for, and the impacts of the open-source Kinect and OpenNI software libraries on the field. The survey's structure is based on a novel hand gesture recognition technique. The literature research revealed 13 strategies for localising hands and 11 ways for categorising motions. Only eight application category types were discovered (and three applications accounted for 18 of the 24 papers that included real-world applications to test a gesture recognition system). The OpenNI hand tracking approach is sufficient for the applications that have been tried thus far, as shown by the publications that use the OpenNI libraries and the Kinect for hand tracking, which generally focus on applications rather than localization and classification techniques. Demanding situations haven't yet been used to test the limits of the Kinect and other depth sensors for gesture recognition. The Microsoft Kinect's video camera (VGA, 640x480) and depth camera (QVGA, 320x240) both create picture streams at a rate of 30 frames per second (fps). The depth camera, developed by PrimeSense1 and included in the ASUS Xtion Pro2, operates via structured light. An infrared (IR) light emitter casts a dense, erratic pattern of dots onto a scene, which an IR camera detects.

These dots are placed at varied intervals based on the object's distance from the sensor. Because the pattern and spacing of the projected dots are known, internal computers can determine the separation between each pixel in the picture by comparing the spacing measured in the IR image to the established reference values. Due to close and far thresholds limiting the sensor's capacity to measure depth, the actual range is around 1.2 to 3.5 metres. Additionally, due of the additional light, the sensor functions poorly in direct sunlight since the IR dot pattern is obscured. Systems and extra depth sensors: Time of flight (ToF) cameras and stereoscopic cameras are the most widely used depth sensor types, aside from structured light cameras like the Kinect and Xtion Pro. ToF cameras may measure the reflected light's phase shift or the time it takes for light to travel from the scene to the sensor and back again in order to calculate the pixel depths. At a high frame rate (50 fps), ToF cameras create low-resolution, accurate depth pictures (144x176). One or more ToF cameras are used in nine of the publications being reviewed. Stereoscopic camera systems acquire two pictures simultaneously from a pair of calibrated video cameras, utilising image registration algorithms to build a disparity map that approximately corresponds to per-pixel depth. In comparison to ToF cameras, stereo systems provide less precise depth pictures and require more processing to solve the image registration problem for each pair of images. However, conventional video cameras may be used to make stereo cameras, and they work well in bright illumination. Additionally, there are stereoscopic camera systems on sale, such as the Point Grey Bumblebee6.

Four of the studies we evaluated make use of stereo camera systems. The advantage of depth cameras over colour cameras may be most noticeable when segmenting hands. It is typical to only utilise a depth threshold to isolate the hands in applications where the user is expected to face the depth camera and hold their hands out in front of themselves for gesturing. The hands are recognised using depth thresholding as the locations between various near and far distance thresholds centred on the predicted centroid of the hand's Z (depth) value, which can be established beforehand and presented to the user or identified as the nearest point in the scene. Setting boundaries on the detected hand's area, which entails lowering the anticipated number of pixels in the blob split by depth thresholding, is an excellent approach to reduce noise susceptibility. Li and Klompmaker make use of this adjustment.

Another modification to depth thresholding is to forecast hand depth based on the placement of other body parts rather than assuming the hand is always the nearest item in the picture. Both Cerlinca and Van den Bergh pinpoint the head using OpenCV7's precise face detection before making an educated estimate as to where the probable hands may be. To determine where the hands were, Fujimura employed body detection based on Karhunen-Loeve Decomposition. When separating bodies from the background in depth pictures, Biswas employed depth histograms rather than hard cutoff values to assure continuous areas and reduce noise. Due to the fact that it captures both temporal and geographical information, hand tracking makes a solid foundation for dynamic gesture detection. Fully articulated 20-point body tracking with nodes for each hand is supported by both the OpenNI framework and the Microsoft Kinect SDK. The Kinect was developed especially for tracking whole-body movements.

The NITE body tracking technology could be more well-liked in the scientific community as a result of its earlier availability. NITE also offers waist-up body tracking and wave gesture-based hand identification. NITE requires a calibration position before body tracking can begin, unlike the official Kinect SDK. Using the NITE body tracking system, Avancini, Bellmore, Chang, Frati, Hassani, Lai, Ramey, and Zafrulla successively track the location of the hands. This review summarises the methods for hand localization and gesture classification found in the research on gesture

DOI: 10.17148/IARJSET.2023.10103

recognition, but it also shows that there are little differences in the actual applications used to evaluate these methods. Applications that push depth sensor capabilities and make use of depth data in difficult situations are still absent.

Shrivastava et al. [7] The author of this study proposes an original and expedient technique for dynamic hand gesture recognition utilising Intel's OpenCV image processing library. Many methods for hand gesture detection using visual analysis, such as hidden Markov models, neural networks, and syntactical analysis, have been proposed (HMM). An HMM is recommended for hand gesture recognition in this research. Training and recognition, feature extraction, and detection and tracking are the three sections of the system. In the first stage, hand detection is done in a less conventional way by utilising the L colour space. The feature extraction approach combines hand orientation and HU invariant moments. Using an average recognition rate of more than 90% for lone hand movements, the Forward algorithm is utilised for recognition, and the Baum-Welch algorithm is used for training with the Left-Right Banded (LRB) topology. Because of the use of OpenCV's built-in capabilities, the system is simple to construct, has a high rate of recognition, and can thus be used for real-time applications. This system integrates two strategies to carry out the detection process—hand localization and extraction: (A) Thresholding: To extract the moving object region from a complicated backdrop, the author might use thresholding to the frame difference to find plausible moving regions. (B) Skin Color Segmentation: Using the colour limitation, skin may be immediately recognised. Skin colour is usually determined using the HSV colour space, however the author discovered that it has a lot of noise, especially in complex backgrounds.

Because L colour space is less prone to noise than HSV and because the component in L colour space represents the pixel component's position between red and green while the component represents between yellow and blue, the author chose to use L colour space since our skin colour is primarily made up of a ratio of red and yellow colours. Due to this, L colour space is preferable to HSV for skin colour separation. Following the detecting process, the hand that was removed is tracked. It is necessary to express colour image data using colour histograms as a probability distribution in order to track coloured objects in video frame sequences. However, the colour distribution histogram in video picture sequences alters dramatically with time. The original mean shift approach was modified to create the CAMSHIFT algorithm, which can dynamically adjust to shifting colour probability distributions. It's important to understand that a gesture recognition system's effectiveness is determined on the characteristics that are retrieved. The following conditions must be met in order for the extracted characteristics to be considered good: They must abide by the following requirements: They must meet the following criteria: they must be easy to compute; they must be invariant under rotation, translation, and reflection; and they cannot be repeated. There are a number of characteristics that may be extracted for hand gesture recognition. It is necessary to express colour image data using colour histograms as a probability distribution in order to track coloured objects in video frame sequences.

However, the colour distribution histogram in video picture sequences alters dramatically with time. The original mean shift approach was modified to create the CAMSHIFT algorithm, which can dynamically adjust to shifting colour probability distributions. It's important to understand that a gesture recognition system's effectiveness is determined on the characteristics that are retrieved. The following conditions must be met in order for the extracted characteristics to be considered good: The following requirements must be satisfied: 1) they must be invariant to rotation, translation, and reflection; 2) they must be straightforward to compute; and 3) they must not be able to be replicated. There are a number of characteristics that may be extracted for hand gesture recognition.

1. Evaluation: Given a set of HMM parameters, the aim is to identify an output observable symbol sequence or vector O. The Forward-Backward strategy is used to address the following problem.
2. Decoding: The challenge is to locate an optimal state sequence that, given a set of HMM inputs, is linked to an observable symbol sequence or vector O. The Viterbi technique is used to resolve the issue below.
3. Training: To generate a vector or sequence of observable symbols with the highest probability.

For our proposed system, real-time gesture recognition performed admirably. The C++ code for the system was produced with the Visual compiler. Ten training videos were given by the author for each of the five exercises. When building each gesture model, the author used LRB topology and 5 states to get a recognition rate of 94.33%. A technique for automatically recognising isolated movements is suggested by this work. The recommended system uses Hu invariant moments and hand motion trajectories, also called hand orientations, as features that are rapid and easy to calculate because OpenCV is used. This strategy prioritises recognition speed above accuracy. Additional study will concentrate on improving tracking techniques to boost this system's recognition rate because it suffers greatly from noise. The lecturer at IIIT Allahabad named Dr. Anupam Agrawal, who served as the inspiration for this paper, is also acknowledged by the author. The author is grateful for Siddharth Rautaray's suggestions, which helped him complete the task.

Lai et al. [8] This work presents a real-time dynamic hand gesture recognition system. The representation of the numerals

DOI: 10.17148/IARJSET.2023.10103

one through nine uses eleven distinct forms of dynamically sensed hand motions. A dynamic video records the moving pictures. The hand shape and skin tone are separated from the busy backdrop by the author using the YCbCr colour space transformation. To identify hand motions, convex defect feature sites on the hand form are identified, together with calculated finger angles and tip locations. The project makes use of OpenCV. Ten test respondents accurately recognised eleven hand signals out of 330 different scenarios, each of which featured three different hand gesture positions. This work presents a real-time dynamic hand gesture recognition system. The representation of the numerals one through nine uses eleven distinct forms of dynamically sensed hand motions. A dynamic video records the moving pictures. The hand shape and skin tone are separated from the busy backdrop by the author using the YCbCr colour space transformation. To identify hand motions, convex defect feature sites on the hand form are identified, together with calculated finger angles and tip locations. The project makes use of OpenCV. Ten test respondents accurately recognised eleven hand signals out of 330 different scenarios, each of which featured three different hand gesture positions. The following gives an example of the three coins algorithm: The three coins were placed on the three dots P0, P1, and P2. The three coins have a red back, a green centre, and a black front. The last coin (the red point) is then positioned before the first coin (the black point), the original rear point is moved to the front, the original front point is changed to the middle, and the original middle point is changed to the back as the three coins are then turned anticlockwise. The original front point is shifted to the middle, the initial median point is erased, and the front is inserted before the middle in the centre of a clockwise rotation. The convex hull of the point set is discovered following the completion of the three coin procedure. To determine the angles between fingers, three convex defect sites (start point, end point, and depth point) must be chosen from hand contour convex flaws. The curvature of the hand characteristics is used to determine the fingertip points.

The positions of the fingertip and the angles between the fingers help differentiate the different hand actions. In this study, eleven hand motions have been estimated and described. Which demonstrates that even in a dynamic image with just one face, one hand, and a difficult background, a hand motion can still be accurately recognised? the 330 samples that the ten test respondents produced for identifying eleven hand gestures, The numerals one through nine are each represented as a triangle with three separate locations (slightly in the centre, slightly to the right, and slightly to the left). The dynamic video defines each hand move. The author utilises stopped movements for between 0.2 and 0.5 seconds to test the efficacy of gesture recognition. If so, processing for gesture recognition will proceed; otherwise, it is skipped. The processing time for each image was around 55 ms, and the accuracy of recognition was above 95.1%. This study makes use of a real-time dynamic hand gesture detection system. The motions are recognised by the fingers' open, straight natural orientations. On an Acer 5750G laptop with an Intel CoreTM i5-2450M 2.5GHz CPU and 4GB of RAM, all experiments were conducted. Win7 OS, Dev-C + +4.9.9.2 with OpenCV 2.1, and C programming are some examples of software. Dynamic gesture recognition is a crucial component of HCI. This research used a real-time dynamic hand gesture recognition system. All hand gestures are recognised by OpenCV and the dynamic video. The three character points of the hand contour's convex defect are described to establish the angles between the fingers, and the placements of the fingertip to calculate the hand movements. Ten test participants provide 330 examples of hand gestures. Eleven hand gestures are used to represent the numbers one through nine. More than 95.1% of the photos could be correctly identified, and processing time for each one was about 55 ms. Future HCI applications like games, robots, robot houses, and so forth may use the real-time dynamic hand gesture detection technology.

Lu et al. [9] Dynamic hand gesture identification is a crucial yet challenging topic in the pattern recognition and computer vision industries. The author of this paper gives a practical method for dynamic hand gesture identification utilising only a Leap Motion controller, along with a unique feature vector that is suitable for recording dynamic hand motions (LMC). These have not been covered by other newspapers. The Hidden Conditional Neural Field (HCNF) classifier receives the feature vector with depth information in order to recognise dynamic hand motions. Two main processes make up the recommended method's systematic structure: feature extraction and classification using the HCNF classifier. The recommended method is evaluated using two dynamic hand gesture datasets with frames captured using an LMC. The Handicraft-Gesture dataset has a recognition accuracy of 95.0% compared to the LeapMotion-Gesture3D dataset's 89.5%. Experimental results show that the suggested method can be used for several dynamic hand gesture recognition tasks. The development of several interactive applications in human-computer interaction has led to a rise in interest in the pattern recognition and computer vision sectors for human action recognition.

Dynamic hand gesture recognition is crucial to understanding human behaviour. Despite this, the work is challenging because to the great degree of form change and substantial finger blockage. The challenge of recording such a wide variety of dynamic hand motions with a monocular video sensor limits the effectiveness of video-based hand gesture detection. The Leap Motion controller (LMC) and Microsoft Kinect sensor, two cutting-edge depth sensors that provide three-dimensional (3-D) depth data of the scene, have significantly improved 3-D hand gesture recognition and object segmentation in recent years. Furthermore, Potter et al. demonstrated that the LMC is capable of recognising hand gestures. Because of this, the author of this study used an LMC to differentiate between dynamic hand gestures. The

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 7.12 ⋇ Vol. 10, Issue 1, January 2023

DOI: 10.17148/IARJSET.2023.10103

LMC generates depth data, which contains details on the orientation of the palm, the locations of the fingertips, the centre of the palm, and other significant regions, in contrast to the Kinect sensor and other depth sensors. Therefore, getting them doesn't need any more computing work. Dynamic hand gesture detection is regarded to be challenging due to sequential modelling and categorization. The approach for classifying depth data frame sequences using the relevant hand gesture model for hand gesture recognition is suggested in this article. The two phases that make up the recommended method's methodical structure are as follows: 2. Following feature extraction, classification using the HCNF classifier. The LMC delivers depth data frames that include rotation, scaling data, finger and hand locations, among other details, unlike the Kinect sensor. As a consequence, feature extraction from the LMC takes less time than from the Kinect sensor. The basis for the features used in this research is laid by the depth data frames' data on palm direction, palm normal, fingers, and palm centre position. The categorization of temporal sequences is carried out via the HCNF Classifier (e.g., speech recognition). An HCNF-based classifier is used in this work to recognise dynamic hand motions. HCNF can take into consideration a variety of elements in addition to the advantages of HCRF. The LeapMotion-Gesture3D dataset and the Handicraft-Gesture dataset are two distinct types of dynamic hand gesture datasets that were generated by the author of Datasets for Dynamic Hand Gestures with an LMC. The depth data frames for each dataset were all gathered using the specific LMC API. LeapMotion-Gesture3D Dataset: At the moment, the MSRGesture3D dataset and the bulk of dynamic hand gesture datasets were captured using the Kinect sensor. In order to compare the effectiveness of this hand gesture recognition approach with other methods, the author built a dataset with an LMC named LeapMotion-Gesture3D that resembles the MSRGesture3D dataset. This dataset contains a subset of ASL gestures. The 12 gestures in the dataset include those for the bathroom, blue, finish, green, hungry, milk, past, pig, and shop. Handicraft-Gesture Dataset: The author developed the Handicraft-Gesture dataset to evaluate this strategy using additional real-world gestures.

The ten movements in this dataset, which were taken from pottery processes, are poke, squeeze, pull, scrape, slap, push, cut, circle, and key tap. Both datasets use a 60 frames per second recording rate for the depth data. Each of the 10 people who provided data for the datasets performed each move three times. Consequently, the LeapMotion-Gesture3D dataset and the Handicraft-Gesture dataset each include 360 and 300 depth sequences of data, respectively. The evaluation criterion in these trials was the mean recognition accuracy of six tests. The remaining three subjects' gestures are used after training with the gestures of the seven subjects who were picked at random for each exam. The remaining three people's hands are employed for testing, and the seven participants who were picked at random are used for training using their motions. Thus, no topic is included in both the test set and the training set. The first set of experiments evaluated the relative significance of different characteristics for the LeapMotion-Gesture3D dataset using the HCNF-based classifier. This study proposes a novel feature vector that is suitable for representing dynamic hand motions. The proposed feature vector, which combines single-finger and double-finger characteristics, has two major benefits. Single-finger features, in the first place, deal with the problem of mislabeling brought on by making dynamic hand motions in diverse places. Second, the double fingers' characteristics can be used to distinguish between different connections between neighbouring fingertips. The HCNF-based classifier takes into account a variety of feature types as well as the complex underlying structure of dynamic hand gesture sequences while attempting to recognise dynamic hand gestures. This method has an accuracy of 89.5% for the LeapMotion-Gesture3D dataset and 95.0% for the Handicraft-Gesture dataset, according to test results. The author has provided an effective technique, not previously described in other papers, for recognising dynamic hand movements utilising only the LMC.

Routray et al. [10] Gesture identification is a crucial problem in computer vision because of its numerous applications in a wide range of fields. Many researchers from a wide range of fields, including computer vision, pattern recognition, and many more, are interested in 3-D gesture detection. The author has little trouble getting past the challenges of seeing motion in loud, dimly lit surroundings. Many issues raised by the traditional method to gesture control in the automobile and other sectors can be resolved by using gesture recognition control using time to flight cameras. A Picozense depth camera is utilised to capture depth frames, and OpenCV is then used for further processing. As a result of PyAutogui's integration, we were able to execute tasks by pointing at the computer. For vision-based applications, selecting the correct camera is essential since noiseless images enhance the performance of the computer vision algorithm. The author used a Picozense DCAM700 camera to detect hand gestures in this system.

The depth camera captures the depth frames, which are subsequently processed using a background script by the SDK's thresholding, grey scaling, contour mapping, and convexity defect spots. This gave the author access to several capabilities, including the ability to threshold video frames and change depth values. They utilised the OpenCV cv2.filter2D() method to flex a kernel with an image. In a 5x5 average, the core component is changed with a newer average price while the 25 pixels below it are added and averaged. The kernel is left on top of the central component. Every single pixel in the image is subjected to this procedure again and again. Photographic Gaussian blurring technique: Instead of a box filter, the Gaussian kernel is employed in this. The task has been completed, cv2.GaussianBlur ().For computations to be correct and odd, the kernel's width and height must coincide. The quality variations in the X and Y

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 7.12 ⋇ Vol. 10, Issue 1, January 2023

DOI: 10.17148/IARJSET.2023.10103

directions, or SigmaX and SigmaY, should be shown individually. If just sigmaX is given, letter Y is considered the same as letter X. The kernel size is utilised to compute the area units if they are all zeros. Gaussian blurring may be used to efficiently remove gaussian noise from a picture. the Haar Cascade Classifier: Utilizing Object Detection for Profit In 2001, Paul Viola and Archangel Jones published a study titled "Rapid Object Detection Employing a Boosted Cascade of Straightforward Features" that introduced the effective object recognition technique known as Haar feature-based cascade classifiers. A cascade performer is trained using a huge number of both positive and negative pictures using a machine learning technique. Eventually, spotting details in several photographs comes naturally. For computations to be correct and odd, the kernel's width and height must coincide. The project OpenCV is focused on detection. OpenCV already has a number of pre-trained classifiers for faces, eyes, smiles, and other features. OpenCV comes include a trainer in addition to a detector. If the hand is known, the Rect (x, y, w, h) ROI for the hand is created in real time when these sites are found, providing the locations of the detected hand if the hand is known. The technology can completely eliminate the necessity for the mouse in such a pandemic condition, ushering in a new age of contactless human-computer interaction (HCI).

III. CONCLUSION

The principle goal of this paper is detect piracy action using IP camera where the movies been recorded cause a huge loss for film industries and leads to unemployement so our idea is to detect the location where the piracy is occurring and prevent it.

REFERENCES

[1] Naumovich, Gleb, and Nasir Memon. "Preventing piracy, reverse engineering, and tampering." Computer, vol. 36, no. 7,pp. 64-71. 2003

[2] Fang, Yikai, Kongqiao Wang, Jian Cheng, and Hanqing Lu. "A real-time hand gesture recognition method." In 2007 IEEE International Conference on Multimedia and Expo, pp. 995-998. IEEE, 2007.

[3] Wagner, Daniel, Gerhard Reitmayr, Alessandro Mulloni, Tom Drummond, and Dieter Schmalstieg. "Real-time detection and tracking for augmented reality on mobile phones." IEEE transactions on visualization and computer graphics, vol. 16, no. 3, pp. 355-368. 2009

[4] Garg, Pragati, Naveen Aggarwal, and Sanjeev Sofat. "Vision based hand gesture recognition." International Journal of Computer and Information Engineering, vol. 3, no. 1, pp. 186-191. 2009

[5] Sudler, Hasshi. "Effectiveness of anti-piracy technology: Finding appropriate solutions for evolving online piracy." Business Horizons, vol. 56, no. 2, pp. 149-157. 2015

[6] Choi, Seul-Ki, and Jin Kwak. "Feature analysis and detection techniques for piracy sites." KSII Transactions on Internet and Information Systems (TIIS), vol. 14, no. 5, pp. 2204- 2220. 2020

[7] T. Kong, F. Sun, H. Liu, Y. Jiang, L. Li and J. Shi, "FoveaBox: Beyound Anchor-Based Object Detection," in IEEE Transactions on Image Processing, vol. 29, pp. 7389-7398, doi: 10.1109/TIP.2020.3002345. 2020

[8] Kim, Eui-Jin, and Jin Kwak. "Intelligent piracy site detection technique with high accuracy." KSII Transactions on Internet and Information Systems (TIIS), vol. 15, no. 1, pp. 285-301. 2021

[9] Pulungan, Ali Basrah, Zhafranul Nafis, and Muhammad Anwar. "Object Detection with A Webcam Using the Python Programming Language." Journal of Applied Engineering and Technological Science (JAETS), vol. 2, no. 2, pp. 103-111. 2021

[10] Aldriwish, Khalid. "A Deep Learning Approach for Malware and Software Piracy Threat Detection." Engineering, Technology & Applied Science Research, vol. 11, no. 6, pp. 7757-7762. 2021