The Seybold Report

ISSN 1533-9211

DOI: 10.5281/zenodo.10547803

SELF-MONITORING SYSTEM FOR VISION BASED APPLICATION USING MACHINE LEARNING ALGORITHMS

Dr. R. NAVEEN KUMAR

Department of Computer Science and Engineering, Brainware University, Kolkata.
Email: rnaveenkumarooty@gmail.com

AMIT KUMAR BHORE

Department of Computational Sciences, Brainware University, Kolkata.

SOURAV SADHUKHAN

Department of Computational Sciences, Brainware University, Kolkata.

Dr. G. MANIVASAGAM

Associate Professor, Department of CS & IT, Jain University, Bangalore.

RUBI SARKAR

Department of Computational Sciences, Brainware University, Kolkata, West Bengal.

Abstract

Automatic face detection is a complex problem which is concerned with the automatic identification of an individual in a digital image. There are many algorithms through which this process can be carried out. However, there are no solutions to detect faces automatically with low resolutions in various application scenarios. We can implement this project's computer vision system to predict whether the screens are near their vision or not. Monitors placed too close or too far away may cause problems that may lead to eyestrain. Viewing distances that are too long can cause you to lean forward and strain to see small text. This can tire the eyes and place stress on the torso because the backrest no longer provides that support. Viewing distances that are too short may cause eyes to work harder to focus (convergence problems) and may require sitting in awkward postures. For instance, the user may tilt their head backwards or push the chair away from the screen, causing you to type with outstretched arms. But there is no alert system for measuring distance automatically from monitor to eye. So in this project, we can design implementation for automatic alerts based on distance based on face recognition. The minimum distance is 0.38 m (l.2fi.) and the maximum distance is 1.02 m (3.3fi.). it can be achieved by using artificial intelligence. We can use a web camera to capture human head positions and separate the background from foreground head positions. Then using image processing techniques to detect face and recognize. Finally, calculate the distance from the monitor to the face via web camera. If the distance is minimum to the pre-defined threshold value means, an alert is automatically generated and intimate to users without using any sensors. And also extends the approach to design the parent-child framework to send alerts at the time of seeing unwanted websites.

Index Term: Artificial Intelligence, Deep Learning, Distance Monitoring, Face Detection, Vision System.

INTRODUCTION

Nowadays everyone uses computers, laptops in many fields. They are very helpful to get the job done easily and quickly. But using these personal computers has many disadvantages. Invented to be portable and wearable, laptops have now completely replaced desktop

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.10547803

computers. It only supports laptops for a limited time. Because it causes spiral pain, headache, etc. When we use the laptop for a long time, the constant staring at the flickering screen may cause eye pain because the distance between the laptop screen and the keyboard is very small.

Eye redness, itching and cloudiness are some of the most common eye problems. In the existing system, provide built-in monitors that save us from vision problems, as well as LED backlit technology. In the current, many research groups have already proposed different systems to find people or faces in images.

The face recognition process in this project followed a two-step approach. First, the image is filtered to only highlight areas where human skin might be present. This filter was developed using basic math functions and sensor image processing. The second step is to take selected areas of skin and remove the darkest and lightest areas of the map. Empirical testing has shown that the areas removed correspond to areas of the face, typically the eyes and eyebrows, the nostrils and the mouth.

The binary skin map and the original image are used together to detect the face in the image. This technique consists in properly caring for areas of the skin so that holes appear in the areas of the face on the eyebrows, eyes, lips or nose. In theory, all other areas of the skin will have little or no scratching and no holes will be made except for desired facial features. However, there is no advanced sensor that detects whether a person is close to the system or not.

Inbuilt Monitor:

In existing system, provide inbuilt monitors to save our vision problems and also LED technology used with backlighting. Inbuilt monitors are nothing but Deformable mirror membrane devices. Deformable mirrors (DM) are mirrors whose surface can be deformed, in order to achieve wave front control and correction of optical aberrations. Deformable mirrors are used in combination with wave front sensors and real-time control systems in adaptive optics. Ray tracing model to find a good mirror curvature that creates the smallest resolvable spot size. In existing several systems designed for the purpose of finding people or faces in images have already been proposed by numerous research groups.

Functioning of Sensors:

The operation of the sensors was based on a two-stage approach. First, the image is filtered to select only areas that may contain human skin. This filter was developed using basic math functions and sensor image processing. The principle of operation of such a wave front sensor is quite simple. Each lens in the device focuses the incoming radiation to a location on the sensor, and the location of that location indicates the orientation of the wave front averaged over the entrance area of the lens.

Map:

The second step is to take the selected areas of the skin and remove the darkest and lightest areas from the MAP (Mean Average Object Detection Accuracy). Object detection is a well-known computer vision problem in which models attempt to locate the correct objects in images and classify those objects into the appropriate classes.

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.10547803

The map is used as a standard measure to analyses the accuracy of an object detection model. Empirical testing has shown that the areas removed correspond to areas of the face, typically the eyes and eyebrows, the nostrils and the mouth. The binary skin map and the original image are used together to detect the face in the image.

Implementation:

Since the introduction of the personal computer and the realization that it causes work-related health problems, many guidelines have been issued regarding the best viewing angles and distances. The permissible distances are too small and the angles too large. The established relationship between viewing angle and viewing distance is generally ignored. Computer work is done at close range. As part of this project, we can implement a visual evaluation system based on integrated webcams for the distance measurement system. We can take pictures of faces and separate the foreground from the background. Face recognition is a computer technology used in various applications to identify human faces in digital images.

Face recognition also refers to the psychological process by which humans locate and interact with faces in a visual scene. Face recognition can be viewed as a special case of object class recognition. Cascade's HAAR algorithm focuses on detecting human faces in front of them. Any changes to the facial features in the database will invalidate the matching process. First, possible areas of the human eye are found by testing all areas of the valley of the grayscale image. The algorithm is then used to generate all possible areas of the face, including the eyebrows, iris, nostrils, and comers of the mouth. Each possible candidate face is normalized to reduce the flash effect caused by uneven lighting and the ripple effect caused by head movements. Each candidate's fit value of is measured by its projection onto the significance.

After several iterations, all the best performing candidates are selected for further validation. In this phase, facial symmetry is measured and the presence of different facial features is checked for each candidate face. And draw the bounding box and also calculate the distance measurement from the webcams. Then extend the framework to implement parental controls to make your kids aware that they are constantly viewing and accessing unwanted websites. HAAR Cascade algorithm focus on the detection of frontal human faces. Any facial feature changes in the database will invalidate the matching process.

Firstly, the possible human eye regions are detected by testing all the valley regions in the gray-level image. Then algorithm is used to generate all the possible face regions which include the eyebrows, the iris, the nostril and the mouth comers. Each possible face candidates is normalized to reduce lightning effect caused due to uneven illumination and the shirring effect due to head movement. The fitness value of each candidate is measured based on its projection on the Eigen-faces. After a number of iterations, all the face candidates with a high fitness value are selected for further verification. At this stage, the face symmetry is measured and the existence of the different facial features is verified for each face candidate. And draw the bounding box and also calculate distance measurement from web cameras.

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.10547803

Image Acquisition:

The first stage of any v1s10n system is the photo acquisition stage. After the photo has been acquired, numerous methods of processing can be carried out to picture binarization preprocessing stage ambitions to section captured image. 1s performed within the for record analysis and it the foreground data from the photo to perform the numerous extraordinary imaginative and prescient tasks required nowadays. However, if the photograph has now not been obtained satisfactorily then the intended obligations won't be possible, in spite of the aid of a few shape of picture enhancement.

The fundamental - dimensional picture is a monochrome (greyscale) picture which has been digitized. Describe image as a -dimensional mild depth characteristic f(x, y) where x and y are spatial coordinates and the fee of at any point (x, y) is proportional to the brightness or grey value of the image at that factor. A digitized image is one where

1. Spatial and gray scale values have been made discrete.
2. Intensity restrained across a frequently spaced grid in x and y directions 3. Intensities sampled to 8 bits (256 values).

In this module, we can capture the face images through web cameras in real time. Capture image can by any type and any size.

Foreground Subtraction:

Background subtraction, also recognized as foreground recognition, is a technique in the fields of picture processing and pc vision in which a picture's foreground is extracted for in addition processing (item recognition and so forth.). Generally, an image's areas of interest are items (humans, cars, text and so forth.) in its foreground.

After the stage of photograph preprocessing (which can also include photo de noising, post processing like morphology and many others.) item localization is needed which may also employ this technique. Background subtraction is a widely used technique for detecting transferring objects in motion pictures from static cameras.

The reason within the approach is that of detecting the moving items from the difference between the cutting-edge frame and a reference body, frequently called "heritage photograph", or "history model". Background subtraction is usually finished if the photograph in query is part of a video flow. In this module we can implement binarization algorithm to separate the foreground from back ground.

Distance Measurement:

Measuring distance between a user and a system using a webcam involves using computer vision techniques to estimate the distance based on the size of the user's face or other objects in the image.

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.10547803

Here is a general approach to measuring distance using a webcam:

1. Calibrate the Camera:

To accurately estimate distances, you need to calibrate the camera by finding the intrinsic and extrinsic parameters of the camera. This can be done using a calibration pattern such as a checkerboard.

2. Detect the Face:

Use a face detection algorithm such as Haar cascades or deep learning-based models to detect the user's face in the image.

3. Estimate the Distance:

Using the size of the face detected in the image and the known size of the face in real life, you can estimate the distance between the user and the camera. This can be done using basic trigonometry or more advanced techniques such as stereo vision. It's important to note that the accuracy of the distance measurement will depend on the accuracy of the camera calibration and the accuracy of the face detection algorithm. Additionally, factors such as lighting conditions and user movement can also affect the accuracy of the distance measurement.

No. of Algorithms:

1. Face detection: HAAR cascade algorithm HAAR cascade is an algorithm that can detect objects in images, irrespective of their scale in image and location. In face detection it uses edges or line detection.
2. Face recognition: CNN algorithm Convolutional Neural Network (CNN) In deep learning, a convolutional neural network (CNN) is a special type of neural network that is designed to process data through multiple layers of arrays. A CNN is well-suited for applications like image recognition and is often used in face recognition software.

Working Process Algorithm of HAAR Cascade

Step 1: Pick a pixel location from the image.

Step 2: Now crop a sub-image with the selected pixel as the center from the source image with the same size as the convolution kernel.

Step 3: Calculate an element-wise product between the values of the kernel and sub- image.

Step 4: Add the result of the image

Step 5: Put the resultant value into the new image at the same place where you picked up the pixel location

Working Process of CNN Algorithm:

It is a method used in artificial neural networks to calculate a gradient that is needed in the calculation of the weights to be used in the network.

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.10547803

Step 1: Randomly initialize the weights and biases. Step 2: Feed the training sample.

Step 3: Propagate the inputs forward; compute the net input and output of each unit in the hidden and output layers.

Step 4: Convolution of the error to the hidden layer.

Step 5: Update weights and biases to reflect the propagated errors.

Training and learning functions are mathematical procedures used to automatically adjust the network's weights and biases.

Step 6: Terminating condition.

Alert Message:

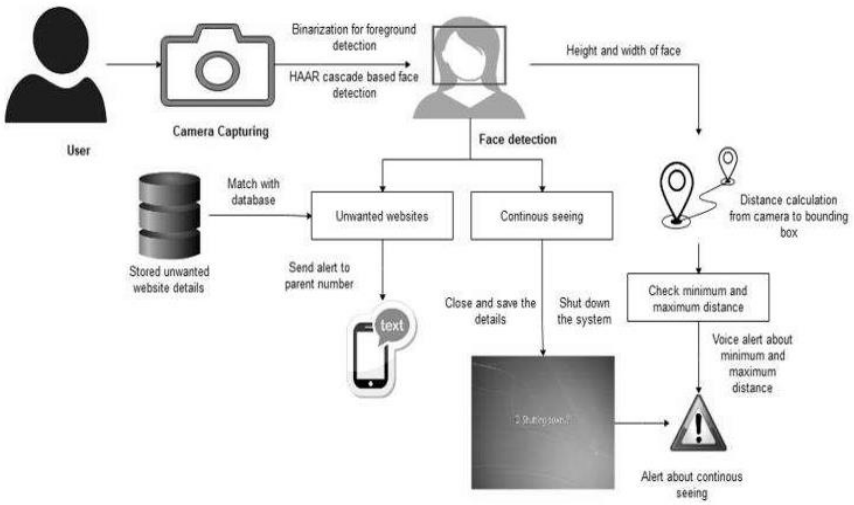
Alert message (or alert notification) is machine-to- person communication that is important or time sensitive. An alert may be a calendar reminder or a notification of a new message. Alert messaging emerged from the study of personal information management (PIM), the science of discovering how people perform certain tasks to acquire, organize, maintain, retrieve and use information relevant to them. In this module implement voice based alert system. We can be set the threshold values for distances. If the distance is less than 1.2 feet means, set voice alert is "Near to the system". If the distance is greater than 3.3 feet means, set voice alert is "Far from the system". And also provide voice alert system for constant seeing the system.

Parental Control:

In this module, provide parental control about children. The system can monitor the constant seeing based on predefined time settings.

After reach the threshold time, alert can be send to stored parents and also mail. Then store the unwanted websites in database for analyzing children can be access the system or not.

Design:



The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.10547803

CONCLUSION

Convergence occurs when the eyes align to the nostrils when we look closely at devices. Convergence allows you to view the image from devices in the same relative proximity on each retina. Without exact coincidence, we see duplicate photos. The closer the devices are, the greater the pressure on the muscles converging toward the eyes. The visual engine also has a resting convergence factor (RPV). It is similar to the shelter's still point, but is an area where the eyes narrow when there may not be an object to focus on. It is also known as dark vergence. It is difficult to set a specific minimum viewing distance limit. If constantly looking closer than the resting convergence point is contributing to eye strain, perhaps we should say that the distance between the eyes should no longer be less than the resting convergence factor. In this activity, we can implement a gadget that uses photo processing techniques to detect faces from a digital camera shot. Then successfully voice faces and place envelopes on face photos. Finally, set distance limits to know if the person is close to the device or not. It also calculated what normal viewing conditions and unwanted websites occur. This device can be beneficial for all seniors in different packages like game apps, business apps etc.

Reference

1) Ruiz, Nataniel, Eunji Chong, and James M. Rehg. "Fine-grained head pose estimation without key points." Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2018.

2) Meyer, Gregory P., et al. "Robust model-based 3d head pose estimation." Proceedings of the IEEE international conference on computer vision. 2015.

3) Cao, Zhiwen, et al. "A vector-based representation to enhance head pose estimation." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021.

4) Yang, Tsun-Yi, et al. "Fsa-net: Leaming fine- grained structure aggregation for head pose estimation from a single image." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.

5) Borghi, Guido, et al. "Face-from-depth for head pose estimation on depth images." IEEE transactions on pattern analysis and machine intelligence 42.3 (2018): 596-609.

6) Drouard, Vincent, et al. "Head pose estimation via probabilistic high-dimensional regression." 2015 IEEE international conference on image processing (ICIP). IEEE, 2015.

7) Kuhnke, Felix, and Jorn Ostermann. "Deep head pose estimation using synthetic images and partial adversarial domain adaption for continuous label spaces." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.

8) Zhang, Hao, et al. "FDN: Feature decoupling network for head pose estimation." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. No. 07. 2020.

9) Wang, Yujia, et al. "A deep coarse-to-fine network for head pose estimation from synthetic data." Pattern Recognition 94 (2019): 196-206.

10) Diaz-Chito, Katerine, Aura Hernandez-Sabate, and Antonio M. Lopez. "A reduced feature set for driver head pose estimation." Applied Soft Computing 45 (2016): 98-107.

11) Borghi, Guido, et al. "Embedded recurrent network for head pose estimation in Ix car." 2017

12) IEEE Intelligent Vehicles Symposium (IV). IEEE, 2017.lx

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.10547803

13) Lathuiliere, Stephane, et al. "Deep mixture of linear inverse regressions applied to head pose estimation." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.

14) Venturelli, Marco, et al. "From depth data to head pose estimation: a Siamese approach." arXiv preprint arXiv:1703.03624 (2017).

15) Xu, Xiang, and Ioannis A. Kakadiaris. "Joint head pose estimation and face alignment framework using global and local CNN features." 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). IEEE, 2017.

16) Diaz-Chito, Katerine, Aura Hemandez-Sabate, and Antonio M. Lopez. "A reduced feature set for driver head pose estimation." Applied Soft Computing 45 (2016): 98-107.

17) Abate, Andrea F., et al. "Near real-time three axis head pose estimation without training." IEEE Access 7 (2019): 64256-64265.

18) Sundararajan, Kalaivani, and Damon L. Woodard. "Head pose estimation in the wild using approximate view manifolds." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2015.

19) Fomalczyk, Krzysztof, and Adam Wojciechowski. "Robust face model based approach to head pose estimation." 2017 Federated Conference on Computer Science and Information Systems (Fed CSIS). IEEE, 2017.

20) Drouard, Vincent, et al. "Robust head-pose estimation based on partially-latent mixture of linear regressions." IEEE Transactions on Image Processing 26.3 (2017): 1428-1440.

21) Khan, Khalil, et al. "Head pose estimation through multi-class face segmentation." 2017 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2017.