Head Motion Detection and tracking with application in Gaming Control

Kartik Patel kartikpatel1995@gmail.com
Manikanta Nallagatla manikanta001nallagatla@gmail.com
Yash Gangrade yashgangrade09@gmail.com

Guided By:
Viraj Shah viraj.shah978@gmail.com
Madhav Pathak madhavkpathak@gmail.com

1 Inception Story

Technology is advancing at a rapid pace, automating many everyday chores in the process. Information technology (IT) is changing the way we perform work and providing society with a multitude of entertainment options. Unfortunately, in the past designers of many software systems have not considered the disabled as active users of technology, and thus this significant part of the world population has often been neglected. A change in this mindset has been emerging in recent years, but slowly every one have started to realize that including this user group is not only profitable, but also beneficial to society as a whole.

One such example of the disability is Alzheimers disease. Alzheimer's Disease is a disease which is mainly caused by stroke. It paralyzes the patients and they looses capability to move any other parts of their body other than their head. The head movement also becomes limited with time and it is the only way they can convey any message. So we basically worked to build a system that converts the head movements to appropriate control actions. Our approach is to track the face and movement of head in 4 directions, i.e., Up, Down, Left and Right. The interaction may not be quite realistic for Alzheimer's patients, but somehow it had to be started at a point. It works with any normal web cam . We have used Open Cv Python as an alternative method to the traditional keyboard input device for tracking head movements to activate keyboard buttons and play Pacman game.

2 Description

Faces are detected using Haar-like features which are the input to the basic classifiers. Object Detection using Haar feature-based cascade classifiers is an effective object detection method proposed by Paul Viola and Michael Jones in their paper, Rapid Object Detection using a Boosted Cascade of Simple Features in 2001[1]. In this approach we will first train the algorithm by providing hundreds of positive and negative images. To extract haar features, images shown in Figure 1 are used. This images are just like convolution kernel. Finally, Algorithm will extract features from data set and save its generalised features in XML file.

Now all possible sizes and location of kernel in 24x24 window has 16000

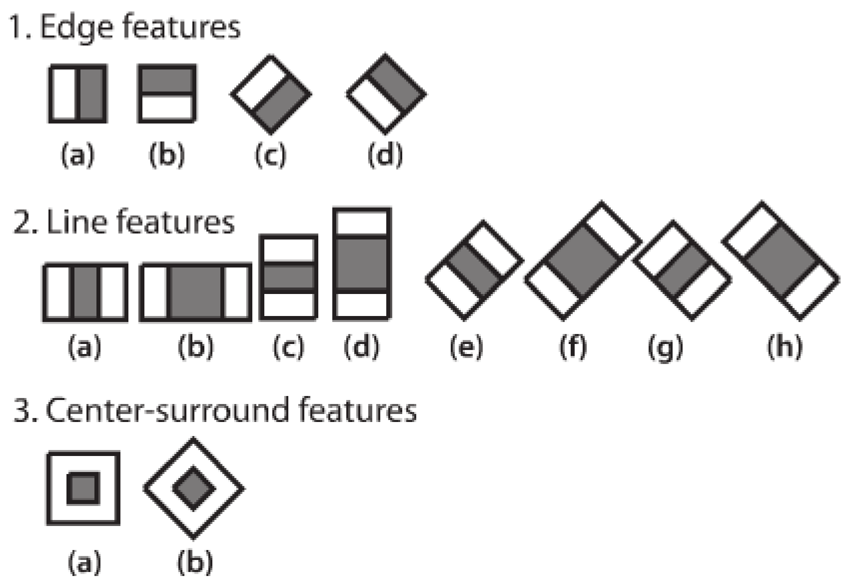

Figure 1: Feature Extraction Figures

features. So, to reduce computation time, instead of applying 6000+ features on a single window, we group the features into different stages. If windows fails any of the stage than it does not contain face. We dont check remaining features.

This is the function which is used to detect face from an image.

```
objects = cv2.CascadeClassifier.detectMultiScale(image[,
          scaleFactor[, minNeighbors[, flags[,
          minSize[, maxSize]]]]])
```

Parameters:

objects Returned rectangular object list

image Image to be detected

scaleFactor Parameter specifying how much the image size is reduced at each image scale.

minNeighbors Parameter specifying how many neighbors each candidate rectangle should have to retain it.

minSize Minimum possible object size. Objects smaller than that are ignored.

maxSize Maximum possible object size. Objects larger than that are ignored.

After detection of the image, our goal was to detect movement of face. For that we have calculated the center point of the rectangle. Also, in every 20 frames we have changed the value of reference point to average centers of detected rectangles. This reference point is used to detect motion for next 20 frames. It was required to avoid need of returning face back to reference point.

Now to detect direction of motion, we have selected a range of rectangle with respect to reference point. If motion is within this rectangle then there is no need of any particular command. This was required to avoid small uncertainty in position. It is also observed that to move head in up direction is much easier than to move down. Hence, downward movement must be more sensitive than upward movement which eventually led to final detection algorithm of movement direction detection.

3 Application

Video Surveillance

Video surveillance systems are now an essential part of daily life. Xie et al. presented a video-based system that captures the face or the head by using a single surveillance camera and performs head tracking methods to detect humans in the video scene.[2]

Car Assistant Systems

Driver's attention recognition plays a significant role in driver assistance systems. It recognizes the state of the driver to prevent driver distraction. The driver's head direction can be an indication of his attention.Liu, Kun [3] used head pose detection to recognize driver's attention. Dongwook Lee[4] introduced a system that detects when the driver gets drowsy based on his head movement.

Accessibility and Assistive Technology

Head movement has been found to be a natural way of interaction. It can be used as an alternative control method and provides accessibility for users when used in human computer interface solutions. Sasou A[5] introduced a wheelchair controlled by an acoustic-based head direction estimation scheme. The user is required to make sounds only by breathing. The direction of the head pose controls the direction of wheelchair. King L. M.[6] also presented a head-movement detection system to provide hands-free control of a power wheelchair. You Song[7] mapped head movement-combined mouth movements to different mouse events, such as move, click, click and drag, etc. People with disabilities who cannot use a traditional keyboard or mouse can benefit from this system.

Eye Tracking Applications

E-Learning

E-learning systems are computer-based teaching systems and are now very common. However, despite the fact that users are usually accustomed to machine interactions, the learning experience can be quite different. In particular, the emotional part is significant in the interaction between teacher and learner and it is missing in computer based learning processes. Calvi C.[8] presented e5Learning which is an e-learning environment that exploits eye data to track user activities, behaviors and emotional or affective states. Two main user states were considered: high workload or non-understanding and tiredness. The author/teacher of the course is able to decide the duration the user should look at certain parts of the course content, whether this content was textual or non-textual. Porta M.[9] sought to build an e-learning platform which determines whether a student is having difficulty understanding some content or is tired or stressed based on the interpreted eye behavior.

Car Assistant Systems

Research is done on applying eye tracking methods in the vehicle industry with the aim of developing monitoring and assisting systems used in cars. For example, an eye tracker could be used in cars to warn drivers when they start getting tired or fall asleep while driving. Driver fatigue can be detected by analyzing blink threshold, eye state (open/closed) and for how long the driver's gaze stays in the same direction. Many eye tracking methods were used in this area of application.

Iris Recognition

Iris recognition is being widely used for biometric authentication. Iris localization is an important and critical step upon which the performance of an iris recognition system depends.

Field of View Estimation

Another interesting application of eye tracking systems is that these systems can serve as an effective tool in optometry to assist in identifying the visual field of any individual, especially identifying blind spots of vision. xiaokun Lee[10] used eye tracking to estimate the field of view to be used for augmented video/image/graphics display.

4 Future Work

Play Full Vision

This report has focused on the situation when one does not have training data of the player whose head movement is being estimated. If there is a desire to get a bit better and interactive performance of the game mode , then one could possibly adapt the player specific information using a relatively small amount of person-specific training data. By this way, computer itself can identify the player and remembers various information about him. Implementing this for various sports such as, Volleyball, Football and Tennis; players may get useful data of their motion, technique and skills for formulating strategic gameplans. One such commercial example is Playfullvision.com

Three Dimensional Modelling

This Head Pose estimation can also be used for three Dimensional modelling of human face and its movements. By this we can narrow the difference between humans and machines which will be useful in Bio informatics, computer graphics, Gaming Industry, Animation and 3-D reconstructions

Mood Estimation

It can be used for Mood estimation of the user. We can guess the mood of the user without need of their involvement by calculating the angle of their sight and comparing the movement of their head with their previous ones.

References

[1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, pp. I–511–I–518 vol.1, 2001.

[2] D. Xie, L. Dang, and R. Tong, "Video based head detection and tracking surveillance system," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*, pp. 2832–2836, May 2012.

[3] K. Liu, Y. Luo, G. Tei, and S. Yang, "Attention recognition of drivers based on head pose estimation," in *Vehicle Power and Propulsion Conference, 2008. VPPC '08. IEEE*, pp. 1–5, Sept 2008.

[4] D. Lee, S. Oh, S. Heo, and M. Hahn, "Drowsy driving detection based on the driver's head movement using infrared sensors," in *Universal Communication, 2008. ISUC '08. Second International Symposium on*, pp. 231–236, Dec 2008.

[5] A. Sasou, "Acoustic head orientation estimation applied to powered wheelchair control," in *Robot Communication and Coordination, 2009. ROBOCOMM '09. Second International Conference on*, pp. 1–6, March 2009.

[6] L. King, H. Nguyen, and P. Taylor, "Hands-free head-movement gesture recognition using artificial neural networks and the magnified gradient function," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pp. 2063–2066, Jan 2005.

[7] Y. Song, Y. Luo, and J. Lin, "Detection of movements of head and mouth to provide computer access for disabled," in *Technologies and Applications of Artificial Intelligence (TAAI), 2011 International Conference on*, pp. 223–226, Nov 2011.

[8] C. Calvi, M. Porta, and D. Sacchi, "e5learning, an e-learning environment based on eye tracking," in *Advanced Learning Technologies, 2008. ICALT '08. Eighth IEEE International Conference on*, pp. 376–380, July 2008.

[9] M. Porta, S. Ricotti, and C. Perez, “Emotional e-learning through eye tracking,” in *Global Engineering Education Conference (EDUCON), 2012 IEEE*, pp. 1–6, April 2012.

[10] X. Li and W. Wee, “An efficient method for eye tracking and eye-gazed fov estimation,” in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pp. 2597–2600, Nov 2009.