Study on Different Approaches for Head Movement Deduction

Vyshagh A, Vishnu K.S

Department of CSE, CALICUT University, Kerala, India

***Abstract*— *Technology, which means the development and application of tools, machines, materials and processes that help to solve human problems, is a wonderful thing. With computers, all the information you could ever want is available on the Internet with a few keystrokes in a search engine. Technology has undoubtedly linked the world together, making it seem so much smaller and interconnected. A Human life getting improves by the use of a recent new application in the field of Human-Computer Interaction (HCI) and Computer Vision (CV). To deal with the versatile use cases is requiring the future desired Perceptual User Interfaces (PUI) and user interface devices for more reliable and fast performance. But this is the scenario for common people like us, physically challenged and mentally disabled are also important in our society. Hence such kinds need special kind of interfaces of interacting with system. The interface we propose is mouse, control by head movements and face gesture. In this survey paper we compare and contrast different algorithm and methods for detecting head movement. We study algorithm like Lucas - Kanade algorithm, Kalman filter algorithm etc. Devices like Gyrosensor, Accelerometer, Memes sensor etc.*

Keywords*— *Adaboost, Camera Mouse, Gyrosensor, MEMS, LK Algorithm.

I. INTRODUCTION

Automatic human detection and tracking is an important and challenging field of research and has many application areas. Monitoring the movements of human being raised the need for tracking. In this survey paper we specify five methods for head tracking. Gyroscopes and Accelerometers can be used in head movement detection systems to obtain information on head movements. A dual axis accelerometer mounted inside a hat was used to collect head movement data. The "Camera Mouse" system has been developed to provide computer access for people with severe disabilities. The system tracks the computer user's movements with a video camera and translates them into the movements of the mouse pointer on the screen. Another popular method is LK algorithm, the Lucas–Kanade method is a widely used differential method for optical flow. It assumes that the flow is essentially constant in local neighborhoods of the pixel under consideration, and solves the basic optical flow equations for all the pixels in that neighborhood. Another model uses MEMS accelerometer to detect the user's head movement in order to direct mouse position on the monitor.

II. SURVEY PAPERS

Design implementation and performance of a cursor control using Gyrosensor introduces several algorithms and electronic components for head tracking and facial gesture. Tracking have been an active research field in the past years as it adds convenience to a variety of applications. It is considered a significant untraditional method of human computer interaction. Head movement detection has also received researcher's attention and interest as it has been found to be a simple and effective interaction method. Both technologies are considered the easiest alternative interface methods. They serve a wide range of severely disabled people who are left with minimal motor abilities.

2.1 Gyro Sensor

Many sensors such as gyroscopes and accelerometers can be used in head movement detection systems to obtain information on head movements. Researches implemented a hands-free head movement classification system which uses pattern recognition techniques with mathematical solutions for enhancement. A Neural Network with the Magnified Gradient Function (MGF) is used. The MGF magnifies the first order derivative of the activation function to increase the rate of convergence and still guarantee convergence, dual axis accelerometer mounted inside a hat was used to collect head movement data. The final data sample is obtained by concatenating the two channels of the dual accelerometer into a single stream to be used in the network with no additional pre-processing of the data. The system needs more experiments to move it from being theoretically proved to being used in real world applications in different scenarios. The accelerometer senses the tilt resulting from the user's head movement. This tilt corresponds to an analog voltage value that can be used to generate control signals. The implementation of this detection method is completely a simulation which represents a good proof of concept. However, it is missing actual accuracy analysis or real-time applicability investigation. The system is very flexible and can be applied in various applications. Gyrosensor method uses relative coordinate system, hence it was found more efficient and accurate, whereas accelerometer uses control signals which involve noise.

2.2 Computer Vision and Accoustic Method

One approach for head movement detection is computer vision-based. A video-based technique for estimating the head pose used it in good image processing application for a real-world problem [1]; attention recognition for drivers. It estimates the relative pose between adjacent views in subsequent video frames. Scale-Invariant Feature Transform (SIFT) descriptors are used in matching the corresponding feature points between two adjacent views. After matching the corresponding feature points, the relative pose angle is found using two-view geometry. With this mathematical solution, which can be applied in the image processing field in general, the x, y, and z coordinates of the head position are determined. The accuracy and performance of the algorithm were not highlighted in the work and thus more work is needed to prove this algorithm to be applicable in real applications.

Some head direction estimation systems localize the user's voice source (i.e. their mouth) to estimate head direction. The method uses a microphone array that can localize the source of the sound. In each head orientation, the localized positions of the sounds generated by the user are distributed around unique areas that can be distinguished. A static head-pose estimation algorithm and a visual 3-D tracking algorithm based on image processing and pattern recognition. The two algorithms are used in a real-time system which estimates the position and orientation of the user's head. This system includes three modules to detect the head; it pro1vides initial estimates of the pose and tracks the head position and orientation continuously in six degrees of freedom. The head detection module uses Adaboost cascades. The initial pose estimation module uses support vector regression (SVR) with localized gradient orientation (LGO) histograms as its input. The tracking module estimates the3-D movement of the head based on an appearance-based particle filter. The algorithm was used to implement a very good driver awareness monitoring application based on head pose estimation. Using this algorithm in real-world applications will be effective for applications similar to a driver awareness monitoring system. It needs more software solution optimization before implementing in real applications. This algorithm can be considered to be a good addition to head tracking systems. The Implementation of head movement tracking system using an IR camera and IR LEDs is done. It tracks a 2 × 2 infrared LED array attached to the back of the head. LED motion is processed using light tracking based on a video analysis technique in which each frame is segmented into regions of interest and the movement of key feature points is tracked between frames. The system was used to control a power wheelchair. The system needs a power supply for the LEDs which could be wired or use batteries. The system can be improved to detect the maximum possible unique head movements to be used in different applications. It needs more experiments to prove its accuracy in addition to more theoretical proofs. The proposition of an effective method that tracks changes in head direction with texture detection in orientation space using low resolution video was costly. The head direction is determined using a LocalBinaryPattern (LBP) to compare between the texture in the current video frame representing the head image and several textures estimated by rotating the head in the previous video frame by some know angles. The method can be used in real-time applications. It can only find the rotation with respect to one axis while the others are fixed. It can determine the rotation with respect to the y axis (left and right rotations) but not neck flexion. Song et al. [6] introduced a head and mouth tracking method using image processing techniques where the face is first detected using an Adaboost algorithm. Then, head movements are detected by analysing the location of the face. Five head motions were defined as the basis of head movements. The geometric centre of the detected face area is calculated and considered to be the head's central coordinates. These coordinates can be analysed over time to trace the motion of the head. The used camera is not head-mounted. The method was found to be fast, which makes it applicable in simple applications for people with disabilities. However, the accuracy and performance of the method were not reported. Presentation a mathematical approach using image processing techniques to trace head movements. They suggested that the pattern of the head movements can be determined by tracing face feature points, such as nostrils, because the movements of the feature point and the movements of the head do not vary widely. Adaboost algorithm was found to be efficient but it has high complexity.

2.3 Camera Mouse

The "Camera Mouse" system M. Betke [2] proposed has been developed to provide computer access for people with severe disabilities. The system tracks the computer user's movements with a video camera and translates them into the movements of the mouse pointer on the screen. Body features such as the tip of the user's nose or finger can be tracked. The visual tracking algorithm is based on cropping an online template of the tracked

ISSN: 2349-6495

feature from the current image frame and testing where this template correlates in the subsequent frame. The location of the highest correlation is interpreted as the new location of the feature in the subsequent frame. Various body features are examined for tracking robustness and user convenience.

The Camera Mouse system currently involves two computers that are linked together—a "vision computer" and a "user computer." The vision computer executes the visual tracking algorithm and sends the position of the tracked feature to the user computer. The user computer interprets the received signals and runs any application software the user wishes to use. The functionalities of the two computers could be integrated into one computer, but the current setup assures sufficient processing power for the visual tracking and allows a supervisor to monitor the tracking performance without interrupting the user's actions.

2.3.1 Vision computer

The vision computer receives and displays a live video of the user sitting in front of the user computer. The video is taken by a camera that is mounted above or below the monitor of the user computer. Watching this video, the user or an attending care provider clicks -with the vision computer's mouse on the feature in the image to be tracked, perhaps the tip of the user's nose. The camera's remote control can be used to initially adjust the pan and tilt angles of the camera and its zoom so that the desired body feature is cantered in the image. The vision system determines the coordinates of the selected feature in the initial image and then computes them automatically in subsequent images.

2.3.2. User Computer

The user computer executes any commercial or custom software application the user chooses. It runs a special driver program in the background that takes the signals received from the vision computer, scales them to, coordinates in the current screen resolution, and then substitutes them for the coordinates of the cursor. The driver program is based on software developed for the Eagle Eyes system and runs independently from the user's chosen application. The Camera Mouse acts as the mouse, and a manual switch box, is used to toggle to the standard mouse and back. The user moves the mouse pointer by moving his or her nose or any other selected feature in space. The driver program contains adjustments for horizontal and vertical "gain" factor. A high gain factor causes small movements of the head to move the mouse pointer greater distances, though with less accuracy. There is a similarity between adjusting the gain and adjusting the camera's zoom.

2.3.3. Tracking Algorithm

When the user initially clicks on the feature to be tracked, a square is drawn around the feature and the sub image within this square is cropped out of the image frame. The cropped sub image is used as a "template" to determine the position of the feature in the next image frame. To find this position, the tracking algorithm uses the template to search for the feature in a "search window" that is centred at the position of the feature in the previous frame. The template is shifted through this search window and correlated with the underlying sub images. The window is defined to contain the centres of all sub images tested.

The experiences with the Camera Mouse system are very encouraging. They show that the Camera Mouse can successfully provide computer access for people with severe disabilities. It is a user-friendly communication device that is especially suitable for children. The system tracks many body features and do not have any user-borne accessories, so it is easily adaptable to serve the special needs of people with various disabilities. To meet the current demand, additional Camera Mouse systems are being installed. A single-computer version of the system is being developed. This method also needs a specific portable camera which is one the disadvantage, although this method is less costly.

2.4 MEMS

MEMS are accelerometer to detect the user's head movement in order to direct mouse position on the monitor. Also clicking events are generated by detecting complete eye blinking as a selection mechanism. The system has been demonstrated to perform mouse movements and clicking events successfully. Here the aim is to build prototype system which is cost-effective, less power hungry, portable and small device substitute for mouse that can help the disabled. Handheld devices such as Personal Digital Assistant (PDA) and smart phones are now widely used for many of our everyday tasks. However, there are at least two reasons that make the interaction on those devices difficult compared to desktop interfaces: small screen size and limited computing power. Pointing and scrolling are one of the most extensively used tasks in almost all computing applications.

The environment in which mobile devices are used is different from that in which desktop is used. The user needs to manage the environment while using the device: holding something, writing notes, opening doors, etc. Also, some users cannot use both hands due to other reasons such as disability or accidents. So, we think that freeing one hand from interaction with the device is very useful. In this work, the use of tilt modality for pointing and scrolling on mobile devices is studied, and evaluated the effectiveness of using tilting in the two tasks then a model for predicting their execution time is developed. The paper describes the implementation of head operated mouse that uses tilt sensors placed on headset to determine head position. Also, it uses 3 axis MEMS accelerometers to detect head tilt in order to direct mouse movement on the monitor. Accelerometer sends the information to the microcontroller. Microcontroller then passes the actual information to encoder. Information encoded is then sent using Transmission to Zigbee receiver. Zigbee receiver will decode the received information. Microcontroller sends to PC through RS232 cable. It will perform the same operation for selecting any documents with the help of eye blink. We constructed an interface system that would allow a similarly paralyzed user to interact with a computer with almost full functional capability. That is, the system operates as a mouse initially, but the user has the ability to toggle in and out of a keyboard mode allowing the entry of text. This is achieved by using the control from a single eye, tracking the position of the pupil for direction, and using blinking as an input. This describes the design of a system that is compatible for all operating systems. Cursor can be moved with the help of head movements. 3-Axis Accelerometer will send the movement direction to Microcontroller. Microcontroller then passes the actual information to encoder. Zigbee receives encoded information and sends to PC through RS232. Also sensor is activated when the user blinks their eye. It allows individuals to operate electronic equipment like communication aids and environmental controls hands free. Each blink of the eye is detected by an infrared sensor, which is mounted on dummy spectacle frames. The eye blink switch can be set up to operate on either eye and may be worn over normal glasses.

The sensitivity of the switch can be adjusted to the user needs and involuntary blinks are ignored. The sensor is connected to a hand-held control unit with a rechargeable battery. This method make use number components which includes zigbee receiver, encoder, decoder etc. which make circuit large. But it is very accurate.

2.5 LK ALGORITHM

Different head movements can help us not only understand accurately whether the speaker in question is in favour of or against the current topic but also recognize human expressions better, i.e. distinguishing similar expressions such as smile and scorn. Since the movements of the feature point -- nostrils on the face and the movements of the head are almost the same, tracing the feature point with Lucas-Kanade(LK) [3] algorithm can find out the pattern of the head movements. GentleBoost classifiers use the difference of coordinates of nostril in an image or a frame to identify the movement of the head, include nodding, shaking, bowing and turning aside. Extract feature points from the area which never move with the change of expressions in human face. We can identify head movement in tracking these points. It was found that there are almost no difference between the images of different people, regardless of their age, sex and race. Pictures in the first row are images of different people, and the images in the second row are their nostrils area. Ultimately, we choose the nostril as the feature point to accomplish our approach. The reason we choose nostril is the most stable and easily recognizable feature point in human face.

This paper present a robust, highly accurate method for head movement recognition based on LK algorithm and GentleBoost. By positioning and tracking the nostril, our method can identify not only nodding and shaking but also bowing and face-turning. When tested on videos captured from different person, the method has achieved average recognition rates not less than 91.4%. In future work we will investigate try to identify the movement for different sizes of faces. Also, we plan to conduct extensive experimental studies using other publicly available face databases. Since LK algorithm revolves around the differential method the complexity is high, but this is very accurate because each pixel in the frame is considered.

2.4.1 Classifier and GentleBoost

In 1990, Schapir proposed an algorithm, which is the original Boosting algorithm and later in 1995, Freund and Schapire [4] have also proposed an improved AdaBoost algorithm. Our method used GentleBoost, an improved AdaBoost algorithm, which is faster than the convergence of AdaBoost, and for better implementation of object detection.

In contrast to AdaBoost, GentleBoost uses real valued features. GentleBoost seams to converge faster than AdaBoost, and performs better for object detection problems. The performance of boosting methods on data which are generated by classes that have a significant overlap, in other words, classification problem where even the Bayes optimal prediction rule has a significant error was occurred. For this case, GentleBoost performs better than AdaBoost since AdaBoost over-emphasizes the atypical examples which eventually results in inferior rules.

III. CONCLUSION

Different approaches for tracking head movement were familiarized. By these approaches physically challenged people can be exposed to the digital world. It was fond that Gyro sensor for head tracking was simple and much accurate when compared to accelerometer. In recent years vibration gyro sensors have found their way into camera-shake detection systems for compact video and still cameras, motion sensing for video games, and vehicle electronic stability control (anti-skid) systems, among other things. Head tracking using MEMS accelerometer was also found to be accurate but still it faces difficulty in cursor movement.lk and gentle boost was found to be robust, highly accurate for head recognition. The method has achieved average recognition rates not less than 91.4%.

REFERENCES

[1] Amer al-rahayfeh and miadfaezipour "Eye Tracking and Head Movement Detection: A State-of-Art Survey" IEEE trans. on Rehabilitation Devices and System,2013

[2] M Betke, J. Gips, and P. Fleming, "The Camera Mouse: VisualTracking of Body Features to Provide Computer Accessfor People with Severe Disabilities," IEEE Trans. on NSRE,Vol. 10, No. 1, pp: 1-10, 2002. I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[3] L. Jian-Zheng and Z. Zheng, ``Head Movement Recognition based on LK algorithm and Gentleboost," in Proc. IEEE 7th Int. Conf. Netw. Comput. Adv. Inf. Manag., Jun. 2011, pp. 232-236.

[4] Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences55(1), 119–139 (1997)