# EyePiano: A Piano Interface for Disabled to Compose Piano Music via Eye Movements

Gimasha Boteju
Dept of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
e-mail: gimabtj@gmail.com

M.U.S Perera
Dept of Computing
Informatics Institute of Technology
Colombo, Sri Lanka
e-mail: udayangi@iit.ac.lk

***Abstract*** **- Human Computer Interaction via eye movements has become a common way of fulfilling disabled people's day to day needs. There is a need of exploring more on computer-mediated performance arts to help disabled people to express themselves in a different manner. Music is the subjected area which disabled people get to involve with this project, as music is a powerful language which can make a great impact on people.**

**As piano is the key instrument in composing music and also most of the people's favourite music instrument, in this paper author present the EyePiano, an alternative to real piano music instrument. Mainly the prototype which has been implemented developed using OpenCV. It was successfully tested and evaluated. Most of the evaluators have rate it as a good project and find it as very useful and effective project.**

## 1. INTRODUCTION

In Human Computer Interaction (HCI) major changes have occurred with computer evolution. From that, Visual-Based HCI is the most widespread area among Audio-Based HCI and Sensor-Based HCI. The main research areas for Visual-Base HCI as follows.

- Facial Expression Analysis
- Body Movement Tracking
- Gesture Recognition
- Gaze Detection (Eyes Movement Tracking)

The exception with eye movement tracking is it's mostly used to help disabled in commanding and action scenario. E.g.: Pointer movement, blinking for clicking.

People with severe motor impairment need this sort of technologies to help them to spend their daily life independently and meaningfully. People who are severely paralysed or afflicted with diseases such as Amyotrophic lateral Sclerosis (ALS) or multiple sclerosis are unable to move or control any parts of their bodies except for their eyes (Chau, *et al.*, 2005).

Though there are products that can be found for people with severe motor impairments, there are limitations as well. One of the main problem is these systems can be expensive. Specially, though eye tracking is powerful alternative for these types of people, when it comes to purchase they are in the range of thousands of US dollars (WebAIM, 2011). Also there is another limitation when it comes to find assistive technologies' compatibility with the operating systems and computer programs (Microsoft, 2011).

Though various techniques are there to help them, more exploration on computer-mediated performance arts is required to express themselves in a different manner.

The Eyewriter is such product which makes people with severe motor impairments to draw using just their eyes (Eyewriter, n.d.). The goal of this research group is to creatively connect patients with ALS to make eye art.

Proposed system is a computer interface for disabled to enable, playing of piano and composition of music via eye movements. In this particular scenario "disabled" refer to the people who have got paralysed or else people having inability to move.

People can get disabled at birth or by an accident. But getting disabled will not make people completely senseless. As they can have ability for some extent and they can be differently able it is required to arrange the suitable environment for them to make the use of their talents or knowledge.

When people get disabled accidently it's a great lost for them both physically and mentally. When they have to make others to do their work it makes them feel powerless. To heal them both physically and mentally it is required to get them involved especially in arts related tasks.

Music is the subjected area which disabled people to get involved with this proposed system. Music is a powerful language which can make great impact on people. It can be a way of exploring the world, simply a

978-1-4673-1666-8/12/$31.00 ©2012 IEEE

way of finding pleasure, a way of relaxing, a way of expressing one's feeling or a medicine to overcome mental and physical problems i.e.: music therapy. Proposed system is a way of enabling disabled people to involve in above set of activities in a real manner.

To make use of proposed system user should be able to move eye balls, blink and hear the sounds. It mainly focuses on the disabled who can be professional composers, musicians, pianists or beginners to learn about music.

## 2. RELATED RESEARCH

When considering about the eye tracking systems for music experiences there are few products currently existing. Basically enabling the severe motor disabled people to play music instruments using eye movements can bring a new experience to their 'locked-in' life. When using eye tracking to play around with music, can implement alternative computer interfaces as replacements for real musical instruments. By providing proper UI and sounds user can get somewhat a similar experience and fun.

### 2.1 Eye Tracking Based Music Systems

EyeGuitar (2010) is an existing product which enables physically disabled to make rhythm based music video games using eye movements. Vickers *et al.* (2010) has initiated this project as a replacement to Guitar Hero which is a hugely popular rhythm based music game. By analysing actual methodology of Guitar Hero they have designed a suitable eye-gaze interaction technique has been designed. Further Vickers *et al.* (2010) states that the test results demonstrated that, users are able to score higher with the gaze technique than using a keyboard for game input. EyeMusic (2008) is a system for composing and performing eye-controlled music and multimedia compositions. To interact with the system user has to move his/her eyes to interact with visual stimuli which triggers a range of musical and visual elements. Though the product mainly does not focus on disabled people it has used eye movements to identify the musical or at least rhythmic of eye movements. Secondly in their research Hornof *et al.* (2008) focuses in giving new opportunities for disabled for musical expression. Polli (2001) design and implement a software system to perform music using eye movements. Based on voluntary and involuntary movements of eyes the software creates a visual and aural landscape. Using high end data transfer and eye tracking technology it concretizes the thought process. Further, according to Vamvakousis and Ramirez (2011) as Polli's composition responded to video images of the eye, not specifically the pupil center.

Oculog (2007) is a system for performing electronic music where a video based eye movement system is used to control the sound. The music exploration is carrying out by vertical and horizontal eye position with blink detection. Kim *et al.* (2007) aimed on users those who cannot hold or touch an instrument. EyeHarp (2010) is a new musical instrument based on eye tracking by Vamvakousis and Ramirez (2010). The instrument allows the user to generate music by controlling sound settings and musical events using eye movements. This method introduces a new way of performing music rather than using hand held instruments. The research mainly let the user compose music than just letting him/her play the instrument.

### 2.2 Eye Tracking Algorithms and Techniques

Commercially well-known eye tracking systems mostly use "corneal reflection/pupil centre" method (Goldberg & Wichansky, 2003). Basically how this kind of system works is when infrared camera is directed to eye, infra red light creates strong reflection. With image processing software locate and identify the features of the eye.

Kreutz (2003) sates that damage to the eye caused by excessive Infrared absorption can happen based on a variety of factors such as time, wavelength, power and the fact that different types of tissue have different absorption rates. It is believed that when dealing with severely motor disabled people it's required to make inventions in order to make their life safe. With use of infrared with corneal reflection method can affect in losing their only communicating way. Therefore it's not a proper approach to follow when it comes to deal with severely disabled people.

Template matching is based on a template which has required type of eye and live video feed from camera, the system will try to locate the user's eye in the subsequent frames with high similarity of that feature. One template (iconic view) for each eye region is available. This will cause severe problems when tracking the eyes because these regions constantly change in appearance (Reinders, 1997). In that case it's required to do required modifications to the general tracking of template matching.

Therefore Chau and Betke (2005) and Reinders (1997) have done enhancements to general template matching in their research in order to overcome its limitations. Chau and Betke (2005) has enhance the algorithms to use it with no loss of accuracy and to perform equally well in all lighting conditions. Reinders (1997) has introduces codebook generation to maintain a set of all possible distinct templates of eye region.

Dispersion-based algorithms, Event detection algorithms, Velocity based algorithms, Hidden Markov algorithm, Minimum spanning tree and Kalman filter are some of the main eye movement classification methods. There are several views by different people on best classification method among them.

Nystrom and Holmqvist (2010) states that Dispersion-based algorithms typically identify gaze samples as belonging to a fixation if the samples are located within a spatially limited region (about 0.5º) for a minimum period of time. The most common dispersion-based algorithm is I-DT, can be found in most of the commercial software. With I-DT algorithm there are limitations such as sensitivity to noise and is poor at providing accurate temporal estimates of event onsets and offsets. Recent work has also shown that the output from I-DT is affected on how dispersion is calculated; keeping other parameters fixed, this factor alone can yield significant differences in the number and durations of fixations (Blignaut, 2009).
According to Komogortsev *et al.* (2010), research results of above 6 types of algorithms on qualitative and quantitative scoring shows as a whole, Kalman Filter method is the best performer for real time eye gaze base interaction. When considering Komogortsev *et al.* (2010) research results they also have not seen considerable performance from Dispersion-based algorithms. Therefore, it is reflected that though Dispersion-based algorithms are mostly used, there are significant amount of limitations with capturing fixations.

## 3. WORK FLOW OF PROTOTYPE

Mainly the system consists of three layers. The top most layer where the user will be interacting with the system and other external devices connecting to the system, consists of interface components.
The main external device which interacts with the system is webcam. The video stream receiving through webcam and Piano UI are the interface components of the system. The second layer consists of two modules where the handling of interface components and integration of two modules happen. As shown one module is to track eyes and other module is for UI which enable the user to interact with system.

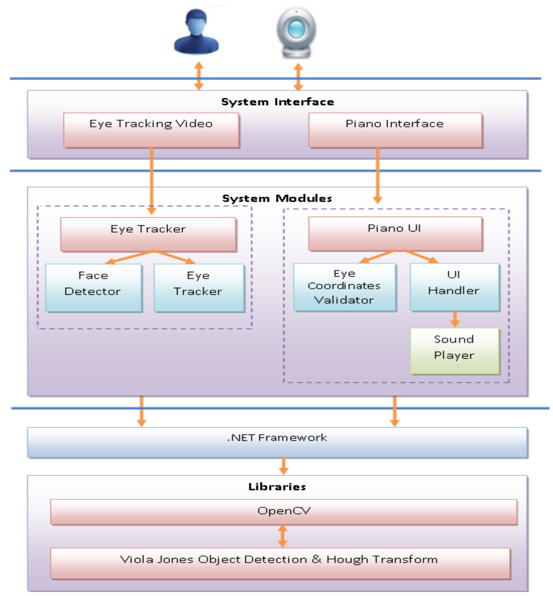


Figure 1: System Architecture

Mainly the eye tracking module consists of two components. First the module start detecting user's face, after that it starts detecting user's pupils. The detecting eye coordinates then will be sending to other module; Piano UI. The functionality of Piano UI module has been divided in to three main components. The UI handler and the sound player do all the UI related functionalities like initialising UI components playing appropriate sound for each piano key and process coordinates receiving from eye tracking module.

Final layer mainly based on .NET Framework, consists of libraries which need to build up the second layer modules. Especially the eye tracking module will be using OpenCV2.1 library to facilitate the tracking of user face and pupils. Technology justification is found later in this chapter.

In order to get the pupil coordinate on UI a simple mapping has been used. This mapping is required as there is no clear cut difference in pupil positions as the eye is a small object. Also, these tracking values which are relative to the capturing frames need to be mapped to the screen coordinates.

Following is the algorithm which was used to facilitate the mapping purpose.

$$screenCoordinatesX = \frac{UIWIDTH * ((pupilCoordinatesX - eyeRegionUpperLeftCornerX))}{eyeRegionWidth}$$

$$\text{screenCoordinatesY} = \text{UIHEIGHT} * \frac{((\text{pupilCoordinatesY} - \text{eyeRegionUpperLeftCornerY}))}{\text{eyeRegionHeight}}$$

## 4. TECHNOLOGIES AND TOOLS

### 4.1 Haar Feature-based Classification

As the main library which has been used for this project implementation is OpenCV face and eye region detection and tracking of face and eye have been done using an algorithm which is available within OpenCV. The algorithm is called Haar Feature-based Cascade Classifier for Object Detection.

This classification mechanism came up with Viola-Jones object detection framework which is the first object detection framework. It allows with a classifier trained with sample views of a particular object to detect this object in a whole image.

When provide a trained cascade of haar classifiers from a file or the classifier database embedded in OpenCV, the algorithm finds rectangular regions which are likely to contain cascade similar objects. Here the cascade refers to a classifier which is trained with hundred sample views of face objects and eye region objects.

These cascade similar objects finding process scans the image several times at different scales. With scanning process some regions overlap with each other which results to be return as sequence of rectangles.

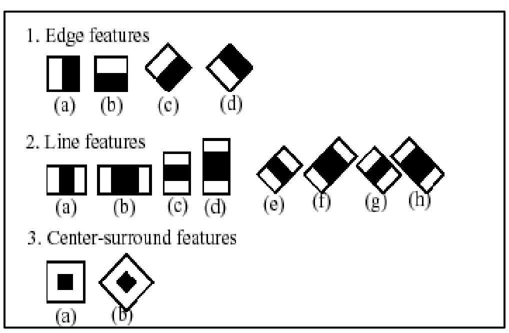


Figure 2: Haar Classifiers

In OpenCV, cvHaarDetectObjects method contains the implementation for this algorithm. cvHaarDetectObjects method does this detection of objects, when pass image and cascade with other image related values as parameters. As the system is real time and it needs faster operation on video stream the settings such as scale_factor = 1.2, min_neighbors = 2, flags =: cmacro: CV_HAAR_DO_CANNY_PRUNING, min_size = *minimum possible face size* (for example, ∼1/4 to 1/16 of the image area in the case of video conferencing) need to be set.

### 4.2 Hough Circles

When tracking pupils with OpenCV one main approach that can follow is detecting circles in a selected region. The Hough Transform is a tool which facilitates identifying objects. Using this technique can isolate features of a particular shape within an image. When figuring out edge pixels using any edge detecting algorithms such as Canny edge detector or Sobel edge detector can get sequence of pixels. With looping through all pixels can figure shape features. This task is difficult as images are never perfect. This lead to having mechanism such as Hough Transform which give more weight age to pixels that are already in the required shape.

cvHoughCircles function is for circle detection with OpenCV. The Circle Hough Transform is a little inefficient at detecting circles, so it uses the gradient method of detecting circles using the Hough Transform (Utkarsh, 2010). But before detecting circles in an image there are some pre-processes need to be carried out on image. Converting image to gray and smoothing it make the results more accurate.

With the frame which needs to find pupils, first need to create the gray image of it with the same dimensions. Secondly, smoothing of image performs so as to prevent a lot of false circle from being detected. Then cvHoughCircles is used to detect circles on the gray image.

## 5. TESTING

### 5.1 Testing Environment

The system was tested using following hardware (basic setup of the system). As the project involves the use of inexpensive web cam, have used Acer laptop integrated web cam.

1. CPU – Intel(R) Core(TM) i3 processor M330 @ 2.13 GHz, FSB 1066 MHz
2. RAM – 4GB
3. MONITOR – 15.6’’ HD LED LCD, 1366 * 768 resolution
4. OS – Windows 7 Home Premium 64-bit
5. Webcam – Acer Crystal Eye Webcam 640 * 480 resolution

Following are the testing scenarios which the system was tested with 8 users, along with the test results obtained.

| Scenario | Test Cases | High Accuracy |
|---|---|---|
| **Webcam position** | <45$^0$, =45$^0$ , >45$^o$ | =45$^o$ |
| **User position** | Different Angles | Center Position 45cm distance |
| **Webcam resolution** | 320 * 240 , 640 * 480 | 640 *480 |
| **Lighting conditions** Natural | 7.30 a.m. , 12.00 noon, 3.30 p.m. , 6.30 p.m. | 7.30 a.m. 12.00 noon |
| Artificial | Table Lamp Halogen Lamp Laptop Screen Light | Table Lamp |
| **Users with glasses** Frames | Full Frame, Frameless | Frameless |
| Lenses | Need to be Tested | - |
| **Different eye colours** | Need to be Tested | - |

Table 1: Testing Scenarios and Results

A regular laptop integrated web cam was used for both implementation and testing phase. Therefore test cases were not selected for different web cams.

The webcam which has used been for this system mainly has two resolution types. They are 640 * 480 and 320 * 240. Higher the resolution becomes the accuracy of results of webcam becomes higher. Therefore, more accurate results were obtained with 640 * 480 resolutions.

The lighting condition is also somewhat a qualitative experiment due to its direct effect on the accuracy and speed of the eye tracking. The accuracy of the system relies heavily on the quality of the captured image. In some cases lighting conditions required is difficult to mention as natural lightning can be different in the same time of different days (sunny or cloudy).

As it is difficult to control the natural lightning conditions to have morning light and daytime light all the time, the system was tested with artificial lighting conditions. Different and inexpensive light sources were used to test the system.

System was tested in order to check whether users with glasses can use the system. With the algorithms which have been used in application problems do not occur when the user is wearing glasses. Same level of accuracy can be expected for a user with glasses and without glasses.

Due to time constraint it was difficult to test with users with different types of lenses and different eye colours. Mainly thick lenses can change the natural look of user's eyes to the external world. Due to different lighting conditions the behaviour of different pupil colours or the features of the pupil can be changed.

It is a constraint of this system that only the user of the system must be there in the capturing video stream.

## 6. EVALUATION

The main concern of carrying out evaluation was to make sure that the prototype meets user requirements and whether it is fitting to the purpose. Therefore, required types of evaluators were selected considering their expertise areas.

All the evaluators found the project idea as interesting and valuable. Regarding technology selection except one person all the other evaluators agreed with the selection of technologies of current prototype which is OpenCV with Viola Jones Haar Cascade method. Almost all the evaluators agreed with the solution that has been introduced with the prototype. Many suggestions were given by the evaluators in order to make the prototype better and to take into next level of the project. Some of the suggested ideas have been added to the prototype and project report after evaluation phase.

Following are some of the future enhancements for the project.

- Improve the accuracy of the prototype by using more robust techniques and algorithms.
- Make the system available online to enable worldwide disabled people to make the use of system.
- Add behaviors of different types of piano to be played by the same set of piano keys of the current system.
- Carry out remaining parts of testing phase.
- Allow multiple key handling at the same time.

## 7. CONCLUSION

With the current system it is difficult to play all the provided set of keys. The accuracy of the prototype need to be further enhanced using the methods such as Kalman Filter along with trained set of data of pupils. As mentioned above some of the test scenarios need to be tested further as they can do a great impact to the accuracy level of the prototype.

## REFERENCES


1. Chau, M. and Betke, M., (2005). *Real Time Eye Tracking and Blink Detection with USB Cameras.* [online] Boston University Computer Science Technical Report. Available From: < http://www.cs.bu.edu/techreports/pdf/2005-012-blink-detection.pdf> [Accessed 28 August 2011]

2. EyeWriter.org, (n.d.). *EyeWriter: About.* [online] Eyewriter.org. Available From: <http://www.eyewriter.org/> [Accessed 09 November 2011]

3. Hornof, A., Rogers, T., Stolet, J., Halverson, T., (2008). *Bringing To Life The Musical Properties Of The Eyes.* [online]. Available From: <http://ix.cs.uoregon.edu/~hornof/downloads/EyeMusic_TechReport.pdf> [Accessed 10 October 2011]

4. Hornof, A. and Sato, L., (2004). *Eyemusic: Making Music With The Eyes.* [online] ACM Transactions on Multimedia Computing. Available from: <http://ix.cs.uoregon.edu/~hornof/downloads/NIME04.pdf> [Accessed 14 Oct. 2011]

5. Kim, J., Schiemer, G., Narushima, T., (2007). *Oculog: Playing With Eye Movements.* [online]. Available From: <http://www.nime.org/2007/proc/nime2007_050.pdf> [Accessed 09 2011]

6. Komogortsev, O. V., Jayarathna, S., Koh, D. H., Gowda, S. M., (2010). *Qualitative and Quantitative Scoring and Evaluation Of The Eye Movement Classification Algorithms.* [online] ACM. Available from: <http://dl.acm.org/citation.cfm?id=1743682&dl=ACM&coll=DL&CFID=48509530&CFTOKEN=84658186> [Accessed on 14 October 2011]

7. Polli, A. (2002). *The Fly's Eye.* [online] The Chicago Tribune Magazine. Available From: <http://www.andreapolli.com/studio/portfolio/eyeportfolio.pdf> [Accessed 09 October 2011]

8. Ramirez, R. and Vamvakousis, Z., (2011). *The EyeHarp: An Eye Tracking-Based Musical Instrument.* [online]. Available From: <http://www.mtg.upf.edu/system/files/publications/Vamvakousis-Zacharias-Master-thesis-2011.pdf> [Accessed 09 October 2011]

9. Urabain, I. R. S. D. (2010). *Eye Tracking For Adaptive Computer Mediated Visual Interaction.* [online] Goldsmiths, University of London. Available From: < http://www.doc.gold.ac.uk/~mas02mb/Cognitive%20Computing/Dissertation/Eye%20tracking.pdf> [Accessed 05 April 2012]

10. Vickers, S., Istance, H., Smalley, M., (2010). *EyeGuitar: Making Rhythm Based Music Video Games Accessible Using Only Eye Movements.* [online]. Available From: <http://www.cse.dmu.ac.uk/~svickers/pdfs/ace2010-vickers.pdf> [Accessed 10 September 2011]


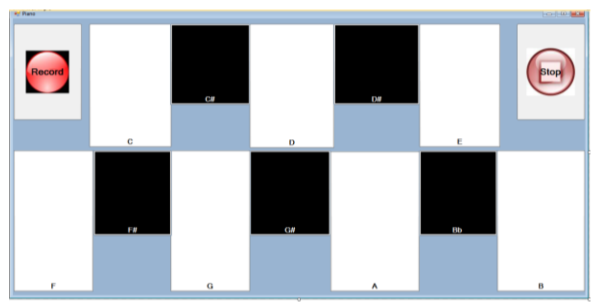


Figure 3: Piano UI

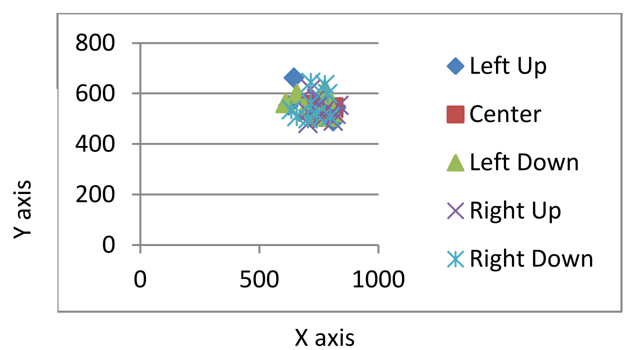


Figure 4: Gaze Placement of UI