EE 367/368 Project Proposal

Real-time Gazed-Controlled Digital Page Turning System

Yao Chen (yaochen1), Zhenzhi Xia (zxia3), Aida Zhumabekova (adazoom)

Introduction

For musicians, turning music sheets by hand while playing an instrument has always posed an unpleasant challenge. The musician often has to stop playing in the middle of the composition to turn the page, therefore a human page turner is commonly being used during a performance. However, during the personal practicing time musicians are left to deal with this problem on their own. A pianist for example, will usually either sacrifice several notes at the end of each page or spend considerable amount of time practicing quickly turning the page so that the music flow is not affected. Although, there is a fair amount of digital page turning applications using bluetooth connected foot pedal or music note detection available on the market, such devices don't provide musicians with enough autonomy because they may require extra foot motion while a pianist needs to pedal the music or, which is the more common downfall, the note detection algorithm will turn the page way too early causing a lot of frustration for musicians. Therefore, we would like to provide a more accessible page turning solution to free musicians' hands and feet and let them devote wholeheartedly to playing music during their practice sessions and even performance. We will do so by providing a software that tracks their eyeball movement simply by using the webcam and turns pages automatically when it receives the indication to do so. We chose the webcam as our medium for this solution because many music players today heavily rely on the electronic music sheets and almost every electronic device has a webcam attached.

Related work

For Human-Computer-Interaction purpose, a page turning system usually consists of a real-time image acquisition system incorporated by the hardware of the electronic device, an eye tracking system, and a feedback system for page turning on software. The core of the page turning system is to track the page turner's eye position e.g. gaze, distance and its orientation in real time. Tremendous work has been done in this field [1-5]. Traditional approach detects object position based on the intensity or color distribution. Eigen-face detection[5] is one of the methods to do so, and it could easily be extended to the eigen-eye detection. However, training such classifier actually requires a large amount of data due to eye shapes, light conditions, etc.. To ease the computational burden of pre-processing the training data, Zhiwei Zhu et al. proposed a methodology for real-time eye tracking under various lighting conditions by combining the bright-pupil effect and the conventional object detection techniques. Morimoto et al. took advantage of pupil-corneal reflection technique for gaze tracking.

Approach

Hardware preparation

- ❏ Use PupilLabs eye tracking camera - a specialized hardware that captures the upclose shot of the eye and detects the pupil. We will only use the camera feed for the purposes of this project and will implement our own pupil segmentation and gaze direction.
- ❏ No use of Android device

Interaction

- ❏ We will specify corner regions on each page which will act as “hot corners” to signal the page turning. The user will then activate the “hot corner” by staring directly at it for a specified short period of time.
- ❏ The user is expected to be sitting in front of the pamphlet with designated “hot corners” that rests on the designated mount, replicating the natural position of a musician reading a music sheet while performing. The distance to the sheet is fixed between 10 - 50 cm, with better accuracy falling around 30 cm range.

Dataset

- ❏ We would use individual frames from the video feed we get with the hardware device and perform image processing techniques we learned in class to detect the pupil and gaze direction.

Implementation

- ❏ We will infer user’s gaze points on a page in real time by using image processing techniques we learned in the class. First, we will decompose the live video feed into frames and apply the following algorithms to locate and track eyes.
- ❏ To extract pupils from within the eye region we will use binarization and thresholding process to detect circles in the image region. Then we detect the center of the pupil by using centroid function. Tracking the position of the centroid will allow us to compute the delta of pupil movement from the video feed stream and recognize when the “hot corner” was triggered.
- ❏ To obtain the gaze direction we will simulate a 3D sphere around the entire eyeball area where the surface of the sphere matches the actual surface of the eye. Then we can use the surface of the simulated sphere to compute the normal vector coming from the surface area corresponding to the pupil. This normal vector will be the gaze direction.
- ❏ We will infer the point of eye gaze intersection with the page by treating the gaze as a ray coming directly from the iris and calculating where it intersects with the plane of the page region within the application window.

Milestones and timeline

Week	Task
1	Set up the live video feed routing a. Use individual frames to establish the working image processing pipeline first
2	Enable pupil detection on the single frame
3	Process the stream of the frames in real time and track user’s pupil movement

4	Add “hot corners” and enable page turning
5	Optimization and debugging

Caveats

- The pupil tracking can possibly be problematic because the natural pupil movement is very quick and chaotic. We will focus on the implementation that relies on the user to maintain the stare at one of the specified “hot corners” for a specified amount of time.
- The real time processing can become another challenge since that would imply considerable time constraints under which we have to produce a reliable output. As a possible solution we can perform our computation on GPU or use parallelization processing to speed up the biggest bottlenecks.

References

[1] Zhu, Zhiwei, Kikuo Fujimura, and Qiang Ji. "Real-time eye detection and tracking under various light conditions." *Proceedings of the 2002 symposium on Eye tracking research & applications*. ACM, 2002

[2] Morimoto, Carlos H., and Marcio RM Mimica. "Eye gaze tracking techniques for interactive applications." Computer vision and image understanding 98.1 (2005): 4-24.

[3] Bellamy, Rachel KE, et al. "Correcting systematic calibration errors in eye tracking data." U.S. Patent No. 9,782,069. 10 Oct. 2017.

[4] Nguyen, Phi Bang, et al. "Method for calibration free gaze tracking using low cost camera." U.S. Patent Application No. 14/389,783.

[5] Turk, Matthew, and Alex Pentland. "Eigenfaces for recognition." *Journal of cognitive neuroscience* 3.1 (1991): 71-86

Goals:
Method:

Plans: