Music Sheets Page Turn with Gaze Detection

**Abstract**

[Replace]

**1. Introduction**

Musical scores are still widely used in this age of internet and electronics. There are solutions to the two-step process: how the performer sends the signal of flipping the page to a device or page-turner, who then performs the action of flipping the page. For pianists, there almost always a page turner by their side to quickly flip pages because of the complexity of piano scores. Violinists or other soloists do not have a page-turner but rather memorizes the score or plan out the intervals while inside a real performance where he or she has a chance to flip the page. This can be extremely hard to achieve for artists whether he is performing or practicing. Therefore, came the need for a substitute. A more available page-turner that can be used at any time. There are several solutions on the market for the process of page turning.

Some products adopted a physical page turn which attaches to a proprietary folder that holds up to forty pages with magnets attached to its plastic page holders. A sweeping arm with magnets attached to the end sweeps from the right to the left, flipping the page. The center of rotation is intentionally placed so that it is not aligned with the center of the two pages, the arm thus automatically detaches from the page leaving it turned.

There are also more restricted solutions of placing a divider such as one invented by PageTurn Inc. between two pages of all books. That saves the trouble of attaching pages to a proprietary folder but sacrifices flexibility on the number of pages that device can turn. It simply moves the divider on a slider to turn pages. This also enables it to turn pages both ways compared to a magnetic arm that only releases pages because of physical distance.

The two physical solutions all come at a high cost averaging more than 100 dollars in price. There are also solutions that simply take advantage of large tablets that can provide a smooth process in turning pages. With Bluetooth, turning pages is much easier to achieve. These devices also come at around half the price of physical turners.

In terms of communication between the page turn device and the performer, almost all took a pedal approach. The two pedals simply navigate pages forward or backwards.

Placing a stationary pedal for signaling page turns limits the performers freedom of movement. An artist’s hands, even feet sometimes, are occupied during performance. Any head or body movement might affect the performance. Facial movements could be too hard to achieve. Eyes are the only highly controllable body part of humans that is not so involved in performance. Thus, comes the need for a reliable method for gaze or eye movement detection.

There are many gaze or eye detection techniques. Yoav Freund and Robert E. Schapire suggested the Adaptive Boost[1] method of training a boosted classifier. It carries the characteristic of possible higher efficiency and lower error rate obtained from the boosted concept of weak learners forming a strong learner by selecting features that tend to improve the predictive power and harden the algorithm. It is used by future studies of feature recognition.

One of those studies is the cascade classifier for object detection with Haar classifiers[1] proposed by Paul Viola and Michael J. Jones in 2001. Different pre-trained Haar classifiers can be used to categorize different objects within the image. Categorizing by ignoring the background and focusing on the important targets improves recognition. However, they have a high error rate, especially when the lighting is different or objects like nostrils stick out inside the image. This makes it extremely easy to be identified. Similarly, a classifier made for irises would likely be easy to make mistakes incorporating a big frame that includes numerous pixel clusters that resembles the desired feature.

Many recent methods use deep learning and neural networks to train more accurate model. The most recent model, YOLOv4, stands out from all others in terms of accuracy and speed, however it does require an enormous amount of calculative power.

Thus, this paper focuses on the process of designing a more balanced page turning solution combined with a cascade classifier-based algorithm. The page turn overcomes the limitation of proprietary folders or sacrifice of page range while still being able to turn pages effectively. The queuing process is replaced with an eye recognition algorithm that is more accurate than cascade classifiers but requires less power than YOLO.

**2. Methods**

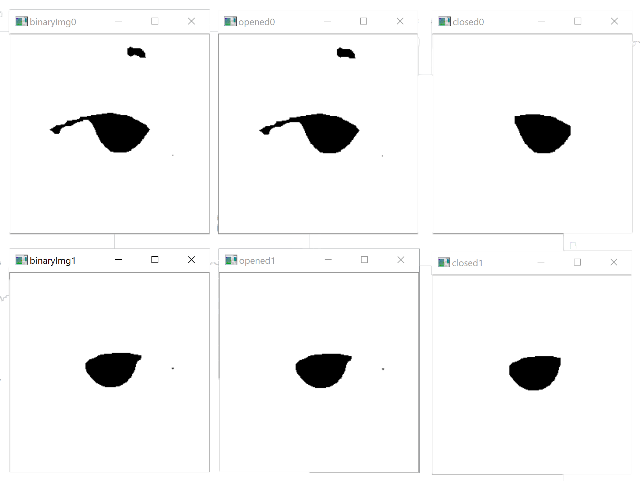
The solution introduced in this paper uses a suction mechanism powered by a pump and a servo to turn pages. The mechanism consists of three parts: a level arm that turns similarly to the magnetic arm solution mentioned above, a suction cup, and a servo.

The suction cup can be used to temporarily anchor on to all kinds of pages when used with a two-way pump. This is unlike the magnetic folder solution since it requires means of attraction from both the page and the page turn mechanism: either a metal-magnet pairing or two magnets. It would also not require a proprietary folder thus preventing any inconvenience without the folder and lessening the requirements of use. Paper is not one-hundred percent air separating. A suction cup might need to be slightly calibrated so that multiple pieces of paper do not get picked up.

The arm is a key component of this mechanism. An arm with a centered point of rotation to the score will result in it not able to turn a second page. The suction mechanism would attach to the paper and turn but it would also stay on the other side of the piece of paper, making changes to the flipped page inevitable. The method of solving this issue is by placing the arm slanted to the left of the center of the opened score. The page to be flipped would still get attached to the arm. However, after the arm almost reached the end of its path it’s position would deviate outside of the score because of the asymmetrical positioning. Thus, after flipping, there will be a clear path for the arm to return to its original position.

A servo provides an accurate method of positioning the arm within a range of 180 degrees. This suffices range requirement of exactly one side of a music stand. It also saves the trouble of adding an encoder in using a DC motor instead. Servos’ disability to be strengthened is a downside of using it to power an arm. The arm that we are planning to use however is not too heavy for the servo to bear. The main frame is designed to be constructed out of balsa wood with a metal horn attached to the end. The overall weight of the arm is not significant enough to be a concern for the servo.

The control of a servo arm is determined with the

 The Open Computer Vision Library(OpenCV) provides many solutions to computer image recognition. The Cascade Classifiers provides a mean…

Opening and closing are two mathematical methods of image processing. Closing

**3. Experimentation**

[R]

[1] <http://www.yorku.ca/gisweb/eats4400/boost.pdf>

[2] <https://www.hpl.hp.com/techreports/Compaq-DEC/CRL-2001-1.pdf>

[3]