Gentry Atkinson, A04115225
CS7321 Take Home Final
Due: 6 December, 2019

Question 1)
Describe 2 eye movement detection algorithms that classify fixations and saccades.

I-MST: eye movement classification based on a Minimum Spanning Tree first applies Prim's algorithm to build an MST across a set of points observed from eye movements of a predefined size. The points are then divided into fixations and saccades by comparison to a predefined distance threshold. The algorithm can be expressed as

Given:
    N, a defined set size
    T, a distance threshold
    G, a graph with vertices V and edges E.
Initialize V to be a set of measured gaze locations such that |V| = N and E = {}
While G is not connected:
    Determine an edge e that is the least distance between to points
    If e does not create a cycle, add e to E

For all v in V:
    If v closest neighbor has distance < T, mark v as Fixation
    Otherwise mark v as Saccade

I-DT: eye movement classification based on Dispersion Threshold defines a temporal window which is translated across the measured data points by shifting one point at a time. A spatial dispersion is calculated for all points in the temporal window which is compared to a defined threshold. If the window encompasses a dispersion that is below the threshold then the first point in the window is marked as a fixation. The algorithm can be expressed as:

Given:
    WS, a temporal window size of at least 100ms
    DT, a dispersion (spatial) threshold

For each point p in a set from p_1 to $p_{size\text{-}WS}$:
    Calculate D=(max(x) - min(x)) - (max(y)-min(y)) of the window from p to p+WS
    If D<DT mark all points from p to p+WS as Fixation
    Otherwise mark p as Saccade

Describe 1 eye movement detection algorithm that can classify fixations, saccades, and pursuits.

I-VDT: eye movement detection by Velocity Dispersion Threshold extends the basic principle of I-DT by including both a dispersion threshold and a velocity threshold. Again, a temporal window is again iterated across the data points. Again, dispersions which fall below the dispersion threshold are classified as Fixations. The remaining points are compared to the velocity threshold. Points below the velocity threshold are marked as Pursuits while points above the threshold are marked as Saccades.

Given:

WS, a temporal window size of at least 100ms
DT, a dispersion (spatial) threshold
VT, a velocity threshold

For each point p in a set from p_1 to $p_{size-WS}$:

Calculate D=(max(x) - min(x)) - (max(y)-min(y)) of the window from p to p+WS
If D<DT mark all points from p to p+WS as Fixation
Otherwise{
Calculate the velocity at p
If velocity < VT mark p as Pursuit
Otherwise mark p as Saccade
}

Discuss the advantages and disadvantages of the above algorithms.

I-MST is an interesting algorithm with a foundation in graph theory but suffers from several substantial disadvantages. First it relies on the collection of a full set of points before it is able to classify any of the points. This will create a lag in time until the full set is collected. Another disadvantage is that it relies on a predefined threshold which must be accurately estimated for the algorithm to perform well. Another is that distance measures between all points in a set can be expensive computationally. Finally it is possible that two fixations will be merged together because the spanning tree is temporarily unaware. The advantage of I-MST is that as a mathematical model it could help us develop a more precise predictive model for eye movements.

I-DT and I-VDT share the same disadvantages of relying on defined thresholds. This disadvantage is doubly true for I-VDT for obvious reasons. They are also particularly sensitive to small saccadic motions (such as an eye adjusting to better fixate on a target) interrupting the temporal window. The advantage is that they are computationally quite lightweight and can easily run on a real-time system.

Question 2)

List 5 technical challenges of an eye-gaze guided system.

- The detection and classification of eye movements need to run in real time without a perceptible lag for the user. The solution to this challenge is partly founded in hardware and partly in software. The hardware can be adjusted so that video oculography is only applied at a low sample rate to confirm the gaze points being generated by a low power approach like photosensors following the motion of a relatively dark pupil. This will generate reliable gaze points without producing a burden of expensive image processing. The software can also be streamlined by employing eye movement detection which is programmatically uncomplicated like IVT of IDT. The classification of pursuits is generally not necessary for an eye guided interface.
- The system will need to distinguish between intentional and unintentional gaze points falling on interactive elements. This challenge can be addressed with good UI design. Making a broad distinction between informative elements (those that will provide the user with information) and interactive elements (those that allow the user to control the system) in terms of location and appearance.
- The system will need to address noise from the user blinking, periodically looking away from the screen, or adjusting the position of their head. This can be addressed with robust eye movement detection such as the Kalman filter.
- The system needs to be comfortable enough not to intrude on the experience of an average user. This means that any approach that requires the head of the user to be rigidly fixed will not be acceptable. Rather a system implemented with a remote webcam or built into a VR headset will be more usable by a human participant.
- The system may not be intuitive for users who are not familiar with the process of eye gaze interaction. The average human is accustomed to gaze being a passive process rather than an interactive one. That is to say that we are not used to objects reacting to us looking at them. This challenge can be overcome with training and intuitive UI design.

Question 3)

List 3 simplifications of the Oculomotor Plant Motor Plant Model.

- Movement about 3 axis is described as being driven by 6 muscles. In reality the motion of the eye is driven by more than a dozen muscles. But the mathematical generality of eye motion being driven by 6 opposed forces produces motions which can closely mimic empirical data.
- Focus is not considered by the model. Gaze is considered to be a line extending from the pupil into infinity rather falling on a single point on that line.
- The fluid momentum of the vitreous humor is not considered. Rather the eye in treated as a solid and non-deformable mass.

Question 4)

Provide a formulation of Fitt's Law:

IP = (ID/MT)

Explain the components

- IP: the index of performance which is a throughput of information / period of time.
- ID: the index of difficulty is calculated as log(2*D/W)
- MT: the movement time from initiation to selection. Equal to a+b*log(2D/W)
- D: the distance from the beginning of a movement to the point of selection
- W: the width of the selection target measured as an intersection of the target and the travel path.

Explain 3 applications of Fitt's Law to HCI research

- The specific throughput of devices can be collected and compared. This gives us a mathematical grounds for comparisons to objectively evaluate the performance of input devices.
- It functions as a predictive model that allows us to design interfaces which will maximize the throughput of a human user.
- It allows us to collect biometric data from several users and compare them. By computing the constants in the MT equation we can make assessments about the health and ability of the user.

Discuss the implications of Fitt's Law to eye gaze controlled interfaces

In order for eye gaze controlled systems to gain popular acceptance in markets or in academia it needs to be demonstrated that they improve the reliability or usability of interfaces which are in common usage such as the mouse or keyboard. This means that we need a method to compare modes of interaction with an interface between many users. Fitt's Law provides a mathematical formality which allows us to compute definite scores as a baseline for comparison between methods of input. Since Fitt's Law is in now way compared to the physical effort of selection or the part of the body used to perform selection we can use it to compare the throughput of any interface that depends on the motion of a cursor and the selection of a target. We can therefore demonstrate the performance of an eye interface in bits/second as compared to a mouse or EEG interface or any other method. In short Fitt's Law provides the common language of comparison for selection activities.