The relationship between input mechanics, flow and auxiliary movements during videogame-play.

Project Report

(Post Mortem)

Autumn 2012

Arelíus Sveinn Arelíusarson
Daníel Sigurðsson
Reynir Örn Björnsson
B.Sc. Computer Science

Instructor: Árni Hermann Reynisson
Examiner: Marta Kristín Lárusdóttir
Professional Advisor: Ólafur Andri Ragnarsson

B.Sc. final project
School of Computer Science

Table of Contents

1. The beginning

As we began our project the first step seemed obvious; we needed to clarify our goals and learn how to achieve them. As we would later discover, our initial visualization of the project was simple and the preliminary version of our work list was naïve. It is evident, when comparing the preliminary work list to the final work list (seen on the next page) that we missed a lot of details. Due to our inexperience there was no way for us to foresee all items right at the start and many items on the final work list emerged along the way. It would also have been useful if right at the start we would have spotted that the project is in fact a three stage project.

Preparations and research
Preparing and conducting the study
Data interpretation, reports and presentations

Although we were aware of this, writing that down and being fully conscious of it would've helped us in planning our scrum sprints.

All of this was however not completely disastrous as the preliminary list gave us an abstract overview of our project and from there we could set up. We intended to create user stories, set up our scrum board and finally, we would divvy up tasks. Then work on the actual project could commence.

2. Setting up our scrum

Before deciding on using scrum, we consulted with teachers and instructors on what methodology would best suit our project. Initially we were leaning towards Kanban. We felt it would provide more flexibility because we knew much of our work was unknown to us. Everyone we consulted recommended scrum rather than Kanban, explaining scrum would keep better track of our time and give us restrictions that would force work to be on time. We accepted these recommendations and began work on our scrum board and creating user stories.

We made a mistake in creating our user stories. First we wrote down broad inexplicit tasks and numbered them. We then broke those stories into smaller tasks noting their relation to the user stories in this manner: If a user story was numbered 11, all tasks it broke down into were numbered 11.1, 11.2, 11.3 and so on [1]. But as we were directly editing our user stories the user stories themselves dissipated. We effectively discarded our user stories and were left with detailed tasks only. In doing so we lost oversight of the project.

3. Work list comparison

Preliminary Work list	***Final Work list***
1 Organizing the project. 2 Finish designing and implementing the game. 3 Find and learn how to use measurement tools. 4 Envision the study. 5 Perform the study. 6 Interpret data and make deliverable.	**1 Project setup** a Declare our null hypotheses. b Set up scrum. c Divvy up responsibilities. **2 Study preparations** a Read up on similar studies. b Learn the structure for deliverable study reports. c Learn methodologies for conducting studies. d Learn about concepts used in the study. e Gather references. **3 Game preparations** a What does a game need to be able to instill flow? b Create user stories. c Needs to display and store data from our questionnaires. d Conduct user studies. **4 Measurement preparations** a Discover what to measure exactly b Learn how to use measurement tools c Learn how to retrieve usable data from measurement tools. **5 Conducting the study** a Considerations b Practice c Apply lessons from Practice **6 Data gathering and handling** a Math. b Query structuring, how do we answer our hypotheses? **7 Deliverables** a Study report b Project report c Code and data?

4. Initial work

We divided up 'fields of work' between us. Daníel programed the game. Reynir set out to discover how we could measure movements and what tools we would use and learn how to use those tools. Arelíus started reading up on similar research to familiarize with the setup of a research report and to discover what has already been done. Moverover, learning about all the concepts we had to deal with was important. What *is* flow? What *exactly* are our questions and *how* do we answer them?

This seemed good to us, there were basically three jobs and three of us and this went, for the most part, quite well.

5. Study preparations

During the study preparations, data was mostly gathered. We needed to know how to present our own study and its goals. We needed to understand and be able to explain concepts within the study. We needed to learn about the methodologies for conducting such studies and we needed to gather references to support our decisions and defend them.

At first our goals were roughly defined. We knew what we wanted to do but we couldn't put it into words. This was fixed shortly after the second status meeting and we have not changed the definition of our goals since then. It was explained to us we needed to set up null hypothesis but our original hypothesis were plain questions.

Reading up on related studies was probably the biggest boon. Simultaneously we gained a lot of understanding on how to conduct such studies and how to configure our study report. Our own study report is using pretty much the same template as one of the studies we read. This reading also gave us a great insight into the concepts we were dealing with as each study methodically explained the concept of flow, its origins and its definition. Single-handedly, this reading was the most precious preparation for the whole project.

The biggest hindrance was finding such studies. We found them, but it took a little work, and often reading a study soon revealed it wasn't appropriate for our own study. We never found a study or directions helping us on how exactly we should measure movement as no study we came across dealt with that particular aspect. At a very late stage, we were told about and given access to a study that was measuring movement, FittsTilt: The Application of Fitts' Law To Tilt-based Interaction. Turns out it does not measure angles of movement but is explaining how to apply Fitts' Law to tilt devices and so was unusable for us.

We also needed to create a questionnaire that wasn't our own fabrication. During this research phase we found that there was a likert scale questionnaire that determined if people were reporting flow. This was a relief because if the scientific community had already agreed upon a list of questions [2] that yield flow reportage we would not have to create and defend each and every question we submitted to our participants.

There is a complete list of all our reading material at the end of this report.

6. Game preparations

As we gained understanding of flow and the variables affecting it our game became more focused. For instance, we discovered that feedback was a crucial component in order to induce flow and so we made displaying score a priority. It was our intention to present questionnaires to participants in-game and let the game keep track of responses and of each participant's' score (to measure success). Originally we planned to write this data directly into a database on the Android device. Once the date for conducting the study approached we realized we did not have time for learning how to implement an SQL database on the Android device, then learn how to insert into it from the device and then how to extract it. We opted for storing answers and score in a text file which we would then device scripts to pump the information into an SQL database. This was a good compromise and enabled us to record responses and score on the device without much hassle. The risk of failing on time by learning, setting up and retrieving an SQL database on Android was much greater than simply learning how to write to a text file with Unity 3d.

During our test runs of the study, we figured there had to be amendments to our game. The game looped through a *preplay questionnaire-gameplay-postplay questionnaire* cycle we needed to prevent participants from starting the next cycle of the loop inadvertently. For this purpose, the night before the study, we added hidden pause buttons at junctures where participants could prematurely start the next phase of the test accidentally. That would be a detriment to our pipeline and force us to restart the process for the current participant. This worked great and we never had to restart a participant because of a participant's mistake.

Also, the night before, we added sound effects. We hadn't given sound effects much priority, but we are very glad we did as they helped enhance the playing experience and thus increase the chances of inducing flow in players.

7. Measurement preparations

The backbone of our study is measuring participants' movements. We had no idea how to do this. We talked with a physiotherapist to garner insight and we discussed the study with the sports department in HR. We had three options for measurement tools; a muscle tensions sensor, an accelerometer and video capturing that would be followed by video analysis with software that measures movement. We wanted two methods employed in the study in case one method would somehow fail. We always wanted the accelerometer as it yielded data that was relatively easy to work with and it wouldn't faze participants as it was just a little device on a headband. We chose to leave the muscle tension sensor behind as we felt it would be too intrusive to ask random people if we could stick a sensor on their skin and it also yielded data we had fewer ideas on how to interpret. Video analysis also seemed easily done. We realized later though, that analysing the videos was very time consuming.

A lot of time went into learning how to use the accelerometer. Data needed to be transformed so we could apply a filter library we intended to use. The filter library, after our configuration, then gave us data we could feed into excel and Unity 3d for calculations.

Learning how the video analysis tool worked didn't take up much time, but once we started testing it we saw that on many occasions you had to watch each video because various effects such as lighting, clothing covering the reference point or something obscuring the scene, caused the software to lose its tracking points. These trials were important; we learned that we had to make clear and concise objects of reference for the software to track. It still did not eliminate the problem altogether. For instance, during the study itself, we had trouble with camera flashes. They would fill frames of the video with a white screen resulting in the loss of tracking references. Sunlight moved across the booth and distorted the tracking references as well. If we would again employ video tracking we would do it in an environment where we could control the lighting. Never being sure if a video would yield undistorted data we had to watch each and every video each time we ran the software on it. Each time a reference point was dropped we had to fix it frame by frame which was excruciating. In the end, we saw that more data was usable from the accelerometer than the video. We reaped the benefits of employing two different measuring techniques as we could ignore the video data and focus on the accelerometer data of which there was more and took less time to prepare.

8. Preparing for the study

We spent a lot of time preparing for the study as we wanted to ensure as much usable data as possible would be gathered within the timeframe we had been allotted. Kringlan had allowed us to do the study in its hallways over a weekend and we were determined to get as much out of it as we could.

The first thing we wanted to do was to envision the entire process from soliciting help from bypassers until a participant leaves us. We were methodical and precise, we noted down every step of the way. We defined roles, assigned duties to roles and then assigned each of us to a role. After that we ran tests to evaluate our proposed method. We asked people from the hallways of our school to help us out. After a participant finished a test run, we noted what went wrong, adjusted our roles and the duties of each role. We amended the game where applicable (adding the invisible pause buttons for instance) and we ran another test. After a few iterations we were happy with our process and had finalized our document detailing our roles, our duties and our precautions.

We are happy with this phase of our project and proud of our methodology and results.

9. Data gathering and interpretation

We had problems with our interpretations. Math was giving us a hard time. We spent four full days discussing ideas, trying to see how we could visually represent our findings and defend them.

We spent time on how exactly we answer our study's questions. We wanted to be able to do several things; to be able to graph a participant's movements, to be able to take a subset of participants and create a graph that represented their movements and to be able to compare graphs. We found ourselves repeatedly rejecting ideas that we initially found plausible, even after we spent time and effort on implementing them.

An example of such wasted time is our Jón Jónson algorithm [5], although fancy, it didn't add anything concrete to our study, it was just eye-candy that displayed a participant's (or a subset of participants) movements.

Another example is our idea of a sine wave [3] to interpret a participants average depth and frequency of auxiliary movements. We intended it to be a visual representation of movements but again, our logic was flawed as shown in the diagrams below.

We should have consulted a mathematician earlier on. It would have saved us a lot of work and frustration.

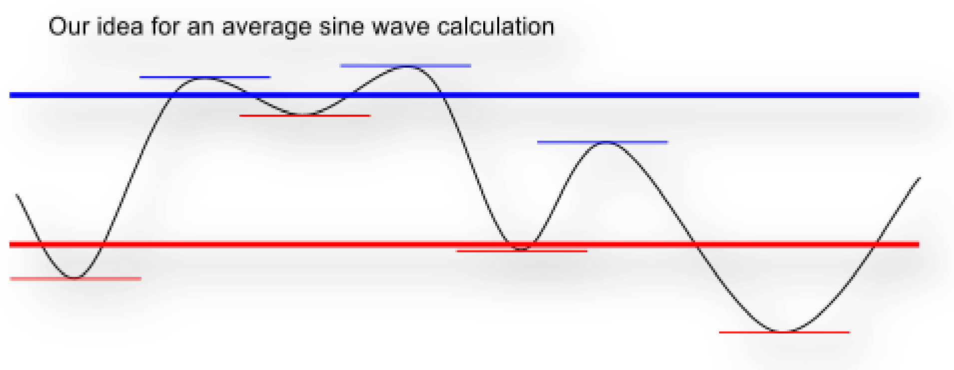

Figure 1: The thick blue line is the average of the tops of a the wave and the thick red line is the average of bottoms, by this we thought we could find out the average depth of a swing to the left and a swing to the right.

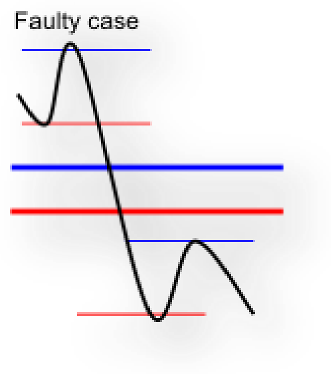

Figure 2: We only discovered this case after work was put into this. As you can see, the thick lines report little to no movement while in actuality there is considerable movement. So we dismissed our sine wave calculations.

10. Presentations

When it came time to present our progress at status meetings we found it irksome that we had in essence no idea how to report it. We didn't really know what was expected of us and answers were fuzzy at best. Our first presentation met with a lot of criticism, our second one was acceptable but our third was terrible. We spent little time on it; we were editing an existing presentation which format didn't really match what we were supposed to say. We didn't rehearse it and our conclusions lacked credibility.

We weren't terribly upset that our delivery was floundering and amateurish, we knew we hadn't rehearsed and we knew this could not let that be the case for our final presentation. So we considered all pointers, jotted them down and were happy to have further guidance in this department for our last stand.

11. Kanban

At later stages in our project, we were finding Scrum to be cumbersome and time consuming to maintain. As we had lost oversight (as explained in the 'Setting up our Scrum' chapter) and the nature of Scrum where a fixed workload is planned over time was a hindrance to us. We needed more flexibility as we frequently had to pause, discard or rearrange jobs. We decided to convert to Kanban to gain that flexibility. We are pleased with this decision, Kanban suited us and the project much better where we weren't tied down by deadlines as extensively and could easily rearrange jobs, pause them and resume them without having to rearrange sprints and without feeling guilty about not meeting deadlines. Our initial instincts had it right all along …

12. Time Management

As we changed from Scrum to Kanban during the course of our project and lost oversight over our Scrum document this resulted in us losing control over our time management. Our document that gives the best impression of the work we put into this project is our Diary. We are however happy to report that all the jobs that needed to be done in the Project Jobs document as well our trello.com Kanban board are completed.

There was no particular job that wasn't finished on time if we discount the creation of our deliverable documents, but they had never been given a due date. We are unhappy about our performance in this area and have learned that time management needs more attention than we gave it.

13. References

[1] Can be seen in the Project Jobs document.
[2] Appendix A.
[3] Appendix B.

14. Reading list

This is a list of what we read to gain insight into studies, concepts and report configurations.

Witmer, Bob G. and Singer, Michael, J. (2011). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3), 225-240. doi:10.1162/105474698565686

Weibel, D., Wissmath, B., Habegger, S., Steiner, Y. and Groner, R. (2008). Playing online games against computer- vs. human-controlled opponents: Effects on presence, flow, and enjoyment. *Computers in Human Behavior*, 24(5), 2274-2291. doi:10.1016/j.chb.2007.11.002

The Flow Scales by Susan A. Jackson, Robert C. Eklund, & Andrew J. Martin (e.d.) Mind Garden, Inc. Retrieved December 11th, 2012, from http://www.mindgarden.com/products/flow.htm

MacKenzie, I. S., & Teather, R. J. (2012). FittsTilt: The Application of Fitts' Law To Tilt-based Interaction. Retrieved December 11th, 2012, from

Weibel, D., Wissmath, B., Groner, R. (2007). Presence vs. Flow in the Context of Computer Games. Retrieved December 12th, 2012 from http://www.temple.edu/ispr/prev_conferences/proceedings/2007/Weibel,%20Wissmath,%20and%20Groner.pdf

Flow (psychology). (n.d.). Retrieved December 12th, 2012 from http://en.wikipedia.org/wiki/Flow_(psychology)

Engeser, S., Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance.
doi:10.1007/s11031-008-9102-4

Nicovich, S. G., Boller, G. W., and Cornwell, T. B. (2005). Experienced presence within computer-mediated communications: Initial explorations on the effects of gender with respect to empathy and immersion. *Journal of Computer-Mediated Communication*, 10(2), article 6.

Witmer, B.G. and Singer, M.J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence, 7(3), 225–240

Madigan, J. (2010). The Psychology of Immersion in Videogames. Retrieved December 12th, 2012 from http://www.psychologyofgames.com/2010/07/the-psychology-of-immersion-in-video-games/

Appendix A - Pre- and postplay questionnaires

This appendix is somewhat in Icelandic because
it is meant to show how we presented our questions to our participants.

An alert before starting a questionnaire:
"Vinsamlegast lesið hverja spurningu vandlega áður en svarað er!"

Preplay Questionnaire:

1. **Aldur**
   A slider let people choose their age.
2. **Kyn**
   KK / KVK
3. **Áttu snjallsíma eða spjaldtölvu?**
   Já / Nei
4. **Að þínu mati, hversu stressuð/aður ertu í dag?**
   Mjög lítið 1 - 7 Mjög mikið
5. **Spilar þú leiki oft?**
   Mjög sjaldan 1 - 7 Mjög oft

Postplay Questionnaire:

Við spilun leiksins þá …

1. … fannst mér áskorunin hæfileg.
2. … voru hugsanir mínar og athafnir fyrirhafnarlausar.
3. … tók ég ekki eftir hvað tímanum leið.
4. … var auðvelt að einbeita sér.
5. … voru hugsanir mínar skýrar.
6. … varð ég djúpt sokkin/n í leikinn.
7. … komu réttar hugsanir og hreyfingar af sjálfu sér.
8. … vissi ég hvað þurfti að gera allan tímann.
9. … fannst mér ég hafa fulla stjórn á leiknum.
10. … varð ég djúpt sokkin/n í eigin hugsanir.
11. … skemmti ég mér innilega.

All questions are answered on a seven point likert scale.

Ósammála			**Hvorki né**			**Sammála**
1	**2**	**3**	**4**	**5**	**6**	**7**

Appendix B - Thoughts about graphing

This appendix illustrates our proposed approaches to graphing and comparing participants' movements. Parts that were used are portrayed and explained better in our Study Report.

Jón Jónsson Graph

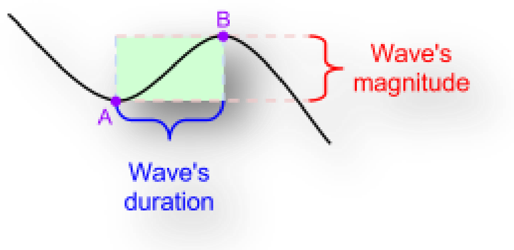

Figure 3

1. Compute column chart from each participant's signal where each column represents a single wave in the signal. I.e. the height of the column represents the magnitude of the wave and the width of the column represents the width of the wave (in other words, how much time the wave spans).
2. Draw a final graph using a sine wave and the column char:

   Initialization:

   a. sineTime = current location in the sine wave (starts at 0)
   b. pointTime = current location in the final graph
   c. interval = interval between points in the final graph
   d. Point.**height** = height of the point currently being drawn into the final graph
   e. Point.**next** = next point to draw
   f. GetColumnAt(**t**) = returns the column at time **t**
   g. Column.**height** = height of a column
   h. Column.**width** = width of a column

   Loop:

   pointTime := pointTime + interval
   Column := GetColumnAt(pointTime)
   sineTime := sineTime + 1 / Column.**width**
   Point.**height** := Sin(sineTime) * Column.**height**
   Point := Point.**next**

Jón Jónsson 2.0

There are 7 steps that need to be explained, those steps are here in bold.

- For each signal
  - Compute **angle** signal from accelerometer signal
  - **Smoothing**
  - Compute average wave **depth**
  - Compute average wave **frequency**
  - Compute **wave chart**
- Compute Jón Jónson Graph from a collection of signals
  - **Merge** all computed wave charts into one
  - Draw a **Sine Wave** according to the wave chart

How to prove that the Jón Jónson graph is legitimate:

- Given signal **A** that starts with a positive direction and always passes 0 before changing directions.
- Given Jón Jónson function **J**
- Then J(A) = A

Details About the Steps

Compute sample angle from sample acceleration

An angle signal is computed like this

- Let G be the accelerometer signal consisting of 3 dimensional vectors.
- Let T be the angle signal
- For each vector $g \in G$ there is a $t \in T$

$$where\ t = 90 - acos(g_y) * \frac{180}{\pi}$$

Smoothing

Each point *p* in the smoothed signal is computed by taking two points, *a* and *b*, on either side of *p* from the original signal, averaging their values and assigning the outcome to *p*.

Each smoothing iteration of the signal has only one parameter; range. The range determines the horizontal interval between *a* and *b*.

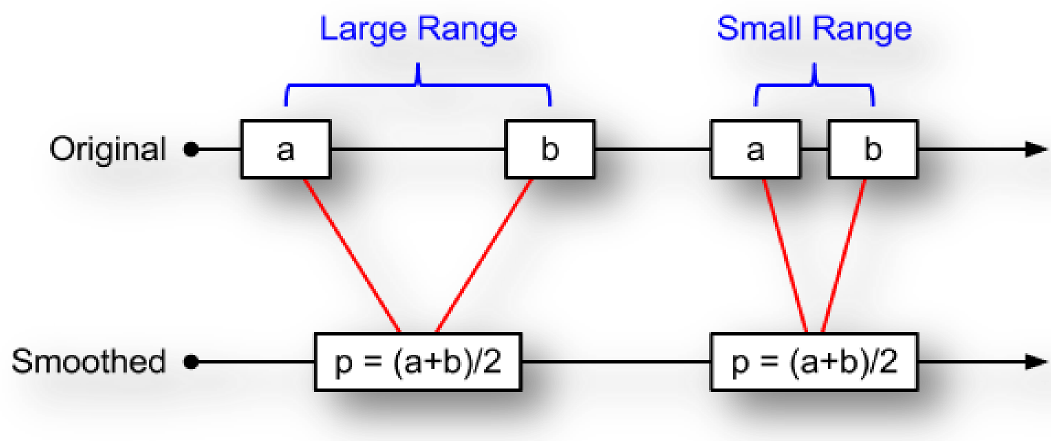

Figure 4

The position of *a* and *b* is half the range away from the position of *p*. If *p* is at position 105 and the range is 2.4 then:

a = [105 - 1.2] = [103.8]
b = [105 + 1.2] = [106.2]

See Figure 4 for details.

This yields a problem at both ends of the signal as shown in Figure 5.

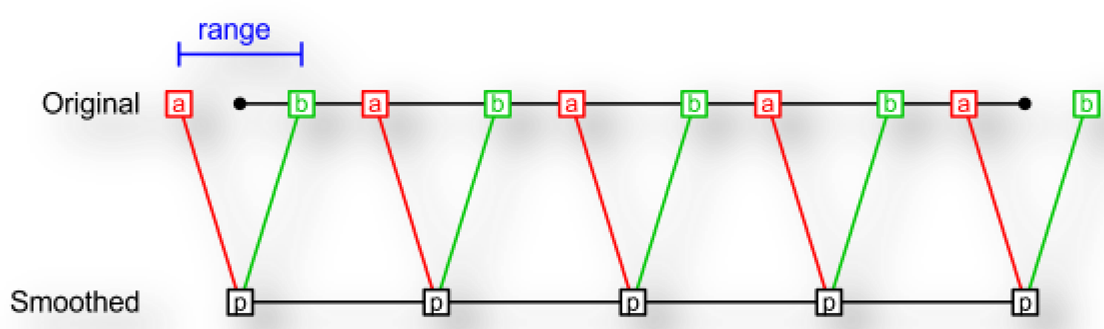

Figure 5

Continuity

To make the original signal seem continuous and not just a list of numbers we simply interpolate between values when reading from the original signal.

If point *a* is at position [14¼], then its value the sum of the following:

75% of the value at index [14]

25% of the value at index [15]

See Figure 6 for details.

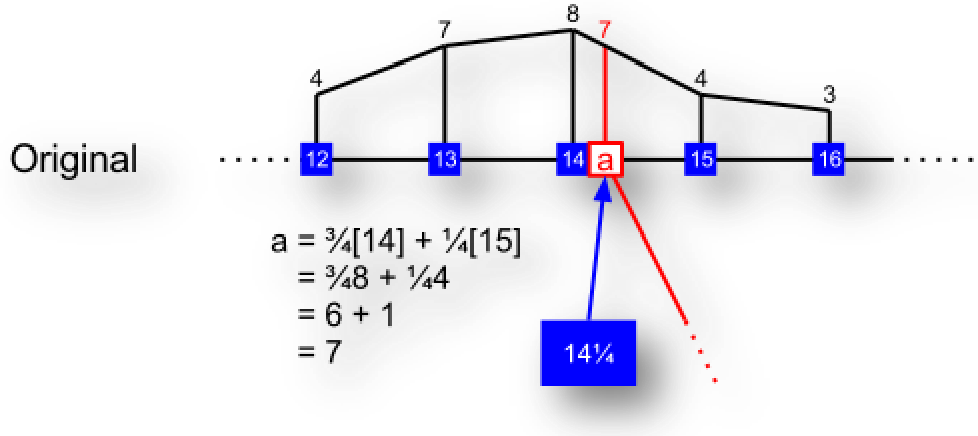

Figure 6

Finity

Since the signal is finite the sample points at the ends of the signal are out of range as seen in Figure 7.

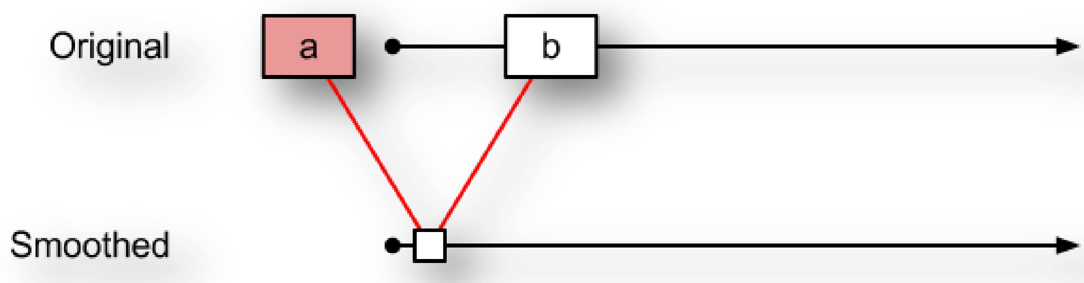

Figure 7

A solution to this problem is to interpolate how much a and b are away from p according to where p is in the signal: So given the following

- n = index of the last point in the signal
- i = index of p
- delta = i / n

Then

- $a = [i - range * delta]$
- $b = [i + range * (1 - delta)]$

See Figure 8 for graphical representation of the solution.

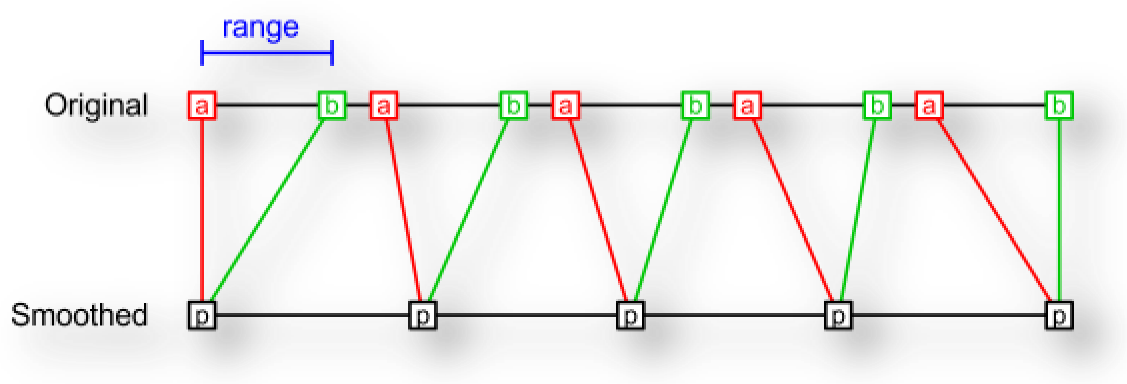

Figure 8

Now all of this smoothing is just one iteration over the signal. The actually smoothing that is applied to each signal is 200 iterations, and the range increments every two iterations, beginning at range 1, and ending at range 100. The range is incremented so that we remove all the frequencies that are too high, not just one frequency. The range 100 was picked because the sample rate of the signal is exactly 640 Hz and thus we would smooth out all frequencies above 6.4Hz (a participant tilting his head both back and forth once every second is doing so at 2 Hz).

This means that a participant would have to shake his head so furiously that he would tilt it both back and forth faster than 3.2 times per second for such motion to be filtered out. Since we're not measuring convulsion we deem it legit to filter out such noise.

Average Sine Wave

Let

- max = the average height of all relative maximums in the signal.
- min = the average height of all the relative minimums in the signal.

Then

- Average Wave Depth = max - min

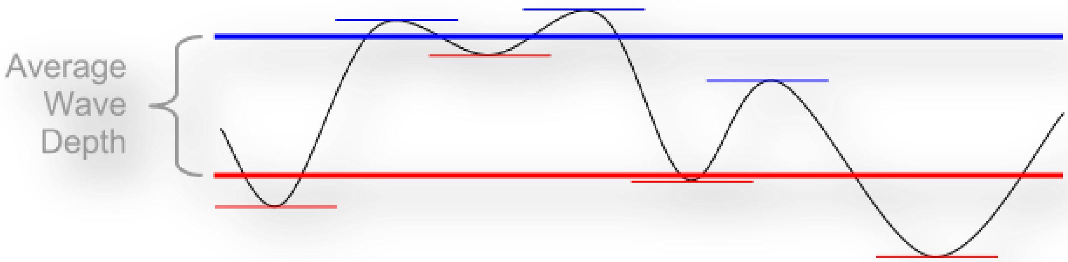

Figure 9: Average Wave Depth

Average Wave Frequency

Simply the number of times the curve changes direction divided by the time span of the whole signal.

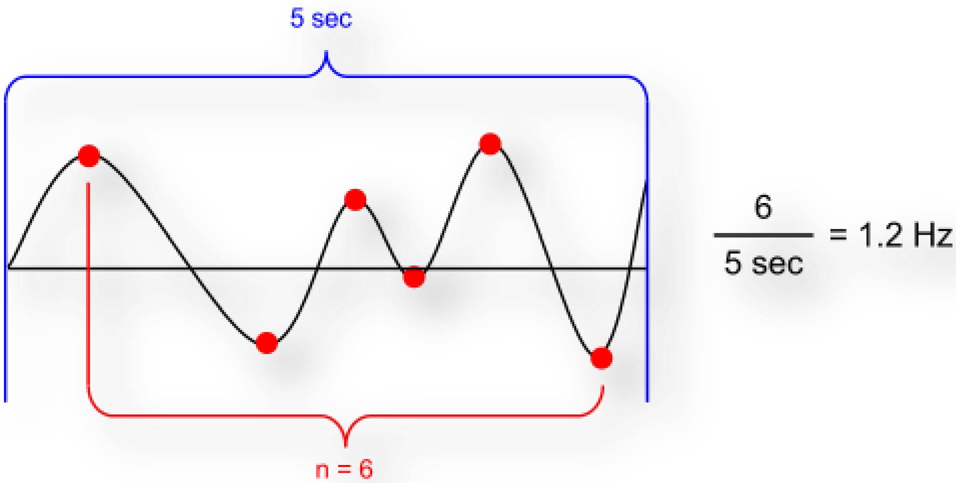

Figure 10: Average Wave Frequency