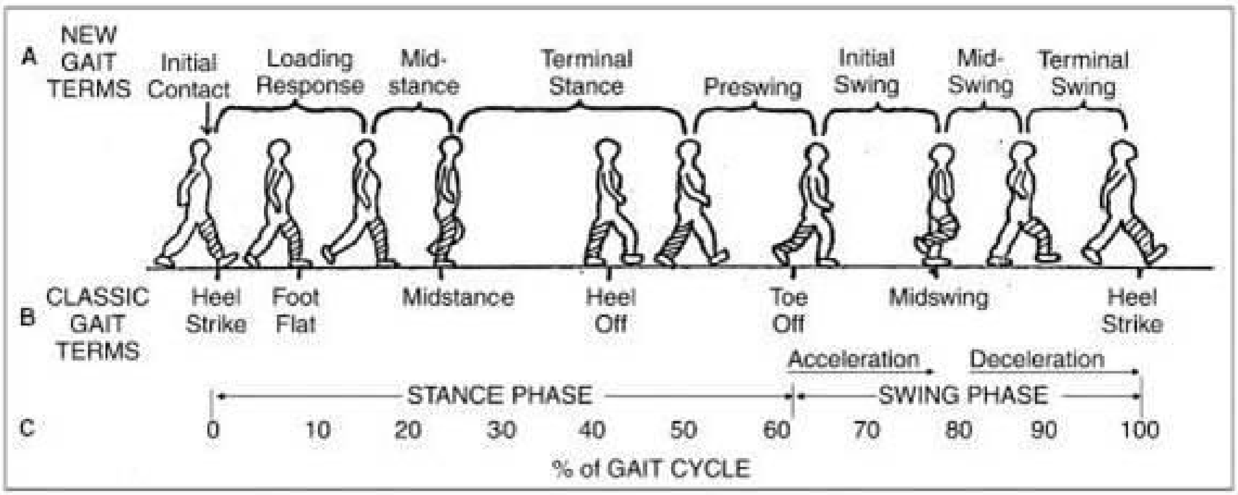
Gait Analysis Report



Nathan Luis, Hunter Mahan, Casey Leong, Quinton Graybeal
Mr. Tronconi
STEM Engineering 2°
October 10, 2018

Abstract

The task for this lab/analysis was to collect sufficient data in order to create a predictive model that could be used to determine a person's height and age from a set of unknown data. The data collected was the acceleration in all three axes, the distance traveled, the time elapsed, and the strides taken. There was also information about each subject that was collected, such as height, leg length, and shoe size. This data was collected to help create the predictive model, as it would be easier to draw conclusions about height with the specific data about each person. The acceleration data was collected through a phone app called Physics Toolbox Accelerometer. This app records the acceleration in all three axes and stores the values recorded in a .csv file which can be uploaded into a spreadsheet for further analysis. To collect the data, the phone was attached to four different subjects in three separate places: the lower abdomen, where the center of gravity would be, the right shin, and the left shin. In these three locations, two trials were taken to improve data accuracy. Each subject would walk ten paces, and the distance, time, and acceleration would be recorded. To help analyze the data and create the predictive model, the average acceleration in each axis of every trial was taken. This, and the time, distance, and steps taken, allowed for the creation of a predictive model which serves as an indicator for height and age. The predictive model based on the data is mostly accurate; however, with something as volatile as a person's gait and unique as a person's body type, the model can drastically change in it's accuracy.

Table of Contents

Introduction

Each individual has a unique way they walk which is called their gait. Ways gait can differ between individuals include time between steps, distance covered per step, the height of the step and the acceleration of the step. The unique combination these characteristics and other similar variables are what makes a person's gait. You can then use gait to determine characteristics of the person in question. Specifically, our goal was to see if we could use a person's gait, and the data collected from a person's gait, to produce a rough estimate of both age and height. We hypothesize that you can predict a person height and age through the following equations.

$$Height\ in\ cm = trunc\left(\frac{distance\ traveled^{(number\ of\ strides/5)}}{average\ length\ of\ stride} * 2.8, 0\right)$$

$$Age\ in\ years = trunc\left(\frac{(time\ elapsed + distance\ traveled)^{(number\ of\ strides/2)}}{2} * \frac{Average\ stride\ rate^{(number\ of\ strides/2)}}{8.5}, 0\right)$$

The height model is based on the assumption that the taller you are, the longer your strides are. This model takes that specific data point and manipulates it further by multiplying it by other data points related to distance and constants to output a number that closely represents the individual's height in cm. The reason the distance traveled is manipulated further and not just multiplied by a constant is because a lot of the numbers are very similar. By putting it to a power, drastically different numbers are achieved. The number produced is also made more different when divided by the average stride length, as this was also only slightly different but when used to divide a bigger number, has better results. This model was created through trial and error, especially with respect to the constant.

The age model is based on the assumption that the older you are, the slower you walk. Similar to the height model, this takes the time elapsed and multiplies it by other time related data points and constants to output a number that is close to the individual's age. The times were all over the place, so in an attempt to make the numbers produce better results, the distance traveled was added to the time elapsed. This was then put to the same power as the height equation and multiplied by a fraction of the stride rate. This model was also created through trial and error. There were many attempts to create a model that would produce a person's age, and this was the closest.

Method

In order to have more reliable data, the four subjects had their name, age, height, center of gravity (from the ground), leg length, arm length, shoe width, shoe length, medical conditions, and if they had insoles or not recorded in a separate spreadsheet. Each subject then had the phone taped to them at their lower abdomen, left shin, and right shin at separate times. They were then instructed to walk 10 steps in a straight line. This walk was timed, and the distance the subject traveled was recorded after they completed their 10 steps. The acceleration was also recorded through a phone app called Physics Toolbox Accelerometer, which takes the acceleration in all three axes and stores the values in a data table. The app started recording at the same time the timing started and the person began to walk. This ensured the data being collected by the app was only data from during the walk. If this were not the case, the average could be affected. This procedure was repeated until there were two trials with the phone in each position.

Results

After performing our measurements of an individual's gait to collect data, we then ran the information through the priorly given equations to calculate their given height and age. The gained information is shown in the chart below.

10 Strides	Location of Phone	Time (s)	Distance (m)	Strides	Avg Velocity (m/s)	Avg Stride Length (m)	Avg A (X)	Avg A (Y)	Avg A (Z)	Avg Stride Rate	Height Predictor (cm)	Age Prediction (Years)
D1a	Center of gravity	7.03	6.63	10	0.943	0.663	0.41	-0.991	0.038	0.703	184	51
D1b	Center of gravity	6.97	6.59	10	0.945	0.659	-0.14	-0.983	0.004	0.697	185	45
D2a	Right shin	7.11	6.62	10	0.931	0.662	0.309	-0.962	0.207	0.711	185	59
D2b	Right shin	7.08	6.65	10	0.939	0.665	0.161	-0.999	0.201	0.708	186	56
D3a	Left shin	6.99	6.6	10	0.944	0.66	0.51	-1.019	0.21	0.699	184	47
D3b	Left shin	7.01	6.58	10	0.939	0.658	-0.026	-1.008	0.189	0.701	184	48
Da1a	Center of gravity	7.53	5.21	10	0.692	0.521	-0.021	-0.994	0	0.753	145	72
Da1b	Center of gravity	7.46	5.23	10	0.701	0.523	0.024	-0.995	0.008	0.746	146	64
Da2a	Right shin	7.39	5.2	10	0.704	0.52	0.142	-0.967	0.232	0.739	145	56
Da2b	Right shin	7.49	5.18	10	0.692	0.518	0.67	-0.984	0.102	0.749	145	66
Da3a	Left shin	7.62	5.22	10	0.685	0.522	-0.253	-0.96	0.06	0.762	146	84
Da3b	Left shin	7.58	5.18	10	0.683	0.518	-0.249	-0.957	0.085	0.758	145	77

N1a	Center of gravity	6.54	6.82	10	1.043	0.682	0.064	-0.991	-0.161	0.654	190	22
N1 b	Center of gravity	6.34	6.79	10	1.071	0.679	-0.077	-0.978	-0.147	0.634	190	15
N2a	Right shin	6.47	6.8	10	1.051	0.68	0.211	-0.988	0.284	0.647	190	19
N2 b	Right shin	6.32	6.83	10	1.081	0.683	0.221	-0.975	0.264	0.632	191	14
N3a	Left shin	6.58	6.81	10	1.035	0.681	-0.244	-0.963	0.228	0.658	190	24
N3 b	Left shin	6.61	6.78	10	1.026	0.678	-0.197	-0.998	0.194	0.661	189	25
C1a	Center of gravity	6.72	6.1	10	0.908	0.61	-0.039	-0.992	0.011	0.672	170	23
C1 b	Center of gravity	6.66	6.12	10	0.919	0.612	0.048	-0.991	-0.068	0.666	171	21
C2a	Right shin	6.59	6.14	10	0.932	0.614	-0.134	-0.998	0.199	0.659	171	18
C2 b	Right shin	6.73	6.08	10	0.903	0.608	-0.167	-0.992	0.184	0.673	170	24
C3a	Left shin	6.79	6.09	10	0.897	0.609	0.19	-0.993	0.153	0.679	170	27
C3 b	Left shin	6.68	6.11	10	0.915	0.611	0.281	-0.96	0.153	0.668	171	22

Analysis

Overall, the data in the charts is shockingly consistent across the boards for a pair of given set of data. For example on the pair of trials for subject Da's left shin where the two sets of data vary .02 m/s when comparing the average velocity. It also has very consistent values for the heights of a given subject. An example from the data is subject C whose calculations for heights is either 170 cm or 171 cm.

The gathered data is not perfect however. This flaw is do to a small sample size and imperfect final calculations. The data set is to small because it only covers a selection of four peoples gait and only for 2 sets of data per areas that the accelerometer was attached. This meant that any major deviation in a given trial could lead to a big departure in the calculation for height and age. We also measured the data with different control variables specifically we only had individuals wear one set of shoes and had different people where different shoes. This could greatly affect the resulting data as certain shoes like high heels could greatly affect the way you are walking. Another flaw in the data is likely the equation used to calculate the age if individuals as it is simultaneously the least accurate at predicting the what it is supposed to and was very imprecise in getting answers.

Conclusion

The predictive models that were created are somewhat accurate, the height model more so than the age model. The problem arises with the assumptions made in the creation of the models. For height, it was assumed a longer distance traveled and average step length meant a taller person. Generally, this is the case, but there are plenty of exceptions. For example, some people of the same height can have different leg lengths, affecting the distance they travel with each step. Even more of a problem is when two people with the same leg length and height take different sized steps. The data collected was not enough to make a very accurate model of someone's height. Despite this, it did come close a lot of the time. Many of the values the predictive model output were within 5 cm of the person's actual height.

The age model is much more temperamental because it is based on the assumption that the slower you walk, the older you are. This is not always the case, similar to the problems with the height model. Some people tend to walk faster than others, and it really varies on a case by case basis. The numbers the model output were all over the place in terms of both accuracy and precision. The values were not close together nor were they close to the actual ages of the subjects, minus a couple outliers.

There are also other sources of error that could have led to the unreliability of the predictive models. Chief among those is the errors in collecting data and the variability in the way people walk. The subjects did not walk the exact same speed or distance every single time, and this led to lots of variations in the data. It created problems with the predictive models because two different sets of data often do not spit out the same number, or close to the same number, when plugged in a formula. There were probably also human errors in the data, whether it was stopping or starting the timer too soon or too late, mismeasuring the distance, recording on the Physics Toolbox application too early or too late, or misentering the collected data into the spreadsheet.

The experiment was overall a success. Two predictive models were created that had some degree of reliability and accuracy with the data collected. Despite many sources of error and perhaps formulas that are not very good, some of the results matched up nicely with the expected values. It is difficult to create something a predictive model that is very precise, especially when the data being dealt with is not the most accurate. There are many things that could be improved upon if this project had to be done again, but there was a lot of learning that happened and growth from not knowing what to do at the beginning to emerging with two semiaccurate predictive models.

References

Tronconi, Claudio. “Gait Analysis.” *STEM Engineering*, Mr. Tronconi, sites.google.com/students.nusd.org/stemse/unit-2/u2-gait-analysis?authuser=2.

Chrystian Vieyra. *Physics Toolbox Accelerometer*. Vieyra, October 5, 2018. Version 1.3.3.

Knight, Randall Dewey, et al. *College Physics: a Strategic Approach*. Pearson, 2018.

Appendix

The following link is to the complete spreadsheets with the entirety of the experimental data and the static data.
https://docs.google.com/spreadsheets/d/1ef8nAzhFUN3W5zNIUHXt_UvIP5oN90_qIv_vTlvYZhQ/edit?usp=sharing