

How much Sample Rate is actually needed? Arm Tracking in Virtual Reality

Hildegardo Noronha
ITI/LARSyS - University of Madeira
Polo Científico e Tecnológico da Madeira,
Caminho da Penteada, piso -2 9020-105 Funchal
Portugal
hildnoronha@gmail.com

Pedro Campos
ITI/LARSyS - University of Madeira
Polo Científico e Tecnológico da Madeira,
Caminho da Penteada, piso -2 9020-105 Funchal
Portugal
pedro.campos.pt@gmail.com

There are plenty of studies dealing with the delays and other relations between head movements and visual response on Virtual Reality setups using head mounted displays. Most of those studies also present some consequences of deviating from those values. Yet, the rest of the human body remains relatively unmapped. In this paper, we present the data found during our research about vision-arm coordination. This data can be used to help build better and more efficient human-computer interfaces, especially those that rely on a virtual avatar with a body and have resource restriction like battery or bandwidth. We tested body tracking Sample Rates ranging from 15 Hz up to 120 Hz (corresponding to total latencies ranging from 37 ms to 95.4 ms) and found out no significant user performance differences. We did, however, find that a small percentage of users are, indeed, capable of noticing the changes in Sample Rate. Based on the found results, we advise that, if one is trying to save battery, bandwidth or processor cycles, a low body tracking Sample Rate could be used with no negative effects on user performance.

Virtual Reality. Arm-Tracking. Sample-Rate.

1. INTRODUCTION

Human body-tracking or motion capture is a field that has boomed with the advent of 3D movies and that is now expanding into gaming. The techniques used for capturing the human body movements and position have been evolving and new techniques keep emerging. Presently, the most commonly used techniques are optical/camera based and inertial based. There are other techniques such as mechanical, magnetic, acoustic and radio reflection tracking.

Even though there are tracking systems and techniques that do not, inherently, have those characteristics (for instance, optical based tracking), those systems can still benefit from this study. The benefits come from the fact that processing power, bandwidth and costs are limited resources and knowing the lower tracking limits of each body part on different situations allows the developer to fine-tune the tracking characteristics as well as the priorities allocated to each body part. This allows the building of cheaper products while maintaining the full quality of the tracking.

2. STATE-OF-THE-ART

The following paragraphs enable us to illustrate the importance of body tracking in the medical field. It also illustrates the importance of a better understanding of the intrinsic values of body tracking for each body part.

(Cloete et al., 2008) compared the kinematic reliability of both inertial and optical motion capture applied to clinical gait analysis. Both systems that were compared were professional, commercially available solutions and were probed at 100 Hz. They found out that the inertial motion capture had more errors than expected but found out that the problem was due to a lycra suit used and that those errors would be solved, based on a paper by (Dejnabadi et al., 2005), if the sensors were secured in place. On the optical side, they encountered issues with markers outside the camera view, shadows, and bad marker reflections. They conclude that the reliability is comparable for lower walking speeds. They also argue that the inertial system is a lot faster to set up than the optical one. This happens because the inertial system is a lycra suit while the optical system is an 8-camera system. The same would not be true if they would compare a strap based inertial system with a 1 or 2 camera system. (Cloete et al., 2010) studied, a couple of years later, the same systems from a repeatability point-of-view and concluded

that inertial systems give enough repeatability to be used on clinical gait analysis. They noted, though, that those systems may perform less optimal on real patients due to body characteristics affecting the sensor placement.

On the studies of the previous paragraph, they used optical tracking systems with 8 cameras. One can argue that it was to achieve a higher degree of accuracy or it may have been due to a lack of better and cheaper solutions at the time (2008 and 2010). In 2012 (Wei et al., 2012) proposed a motion capture method using a single depth camera and compared it with Microsoft Kinect (2012 version; the original version came out in 2010), which is also a single depth camera, and concluded that their method was more accurate.

(Lorincz et al., 2009) ran into a problem that could have been mitigated by the results on this paper. They run a group of sensors on patients, some of them were capturing movement through inertial sensors. Those sensors were fed by a battery and must run up to 18 hours per day. They also ran into issues with data storage and network bandwidth. The high volume of data (reported as 1200 byte/sec/node) as well as a big battery drain might come from the fact that the sensors are set at 100 Hz all the time when they could have been fine tuned to lower values while achieving a similar quality of results. The sample frequency could have been further lowered considering that the movements are not to be interpreted by the user in real time. The authors did throttle down the sensors when battery life was low rising the expected time of battery up to 32h, adding to the importance of more efficient sensor tuning.

(Witchel et al., 2012) made a comparison of four technologies applied to micro-movements. From the technologies they used, the most relevant for this paper, are the 8-camera optical tracking (a Vicon) and an accelerometer mounted on the head. They found out a good correlation between both systems, except for the yaw on the accelerometer. This happens because accelerometers cannot, directly and accurately, measure yaw movements. An important find is that, even without a gyroscope, they were able to match the rotation on the head to an expensive 8 camera tracking system, proving the quality and accuracy of a (striped down, accelerometer only) inertial tracking system.

(Aylward et al., 2007) gives us an example of implementation of inertial sensors on other fields. In this case, the paper focuses on dancing, but it is also tested on baseball illustrating the potential for the tracking of high speed, high acceleration movements while maintaining accuracy.

3. MATERIALS AND METHODS

3.1. Subjects

The experiment was conducted using 41 volunteers (19 to 37 years old). All of them had good knowledge and contact with, at least, one of the following: computers, entertainment systems and gaming systems.

3.2. Materials

3.2.1. Hardware

We used the MPU-9150, a 9-degrees of freedom IMU. The accelerometer has a selectable range from 2 to 16 g. The gyroscope has a selectable range from 250 to 2000 °. The magnetometer has a fixed range of 1200 μ T.

3.2.2. Questionnaire

The users were requested to fill up a questionnaire with following questions graded from -5 to +5:

- (i) I felt that this iteration was -5 – Slower; 0 – Equal; 5 – Faster.
- (ii) I felt that this iteration was -5 – Less Responsive; 0 – Equal; 5 – More Responsive.
- (iii) I felt that this iteration was -5 – Harder; 0 – Equal; 5 – Easier.

3.2.3. Inertial Tracking

The system uses a series of inertial sensors positioned on several bones (arm, forearm, and hand – see Image 1). After some research and trial-and-error, we ended up using (Madgwick et al., 2011)'s algorithm to fuse the sensors' data.

3.3. Experimental Setup

3.3.1. Arm-Tracking

The sensors were placed in the arms so we can know the current orientation of each part of the arm. We then make use of forward kinematics to calculate the exact position of each bone.

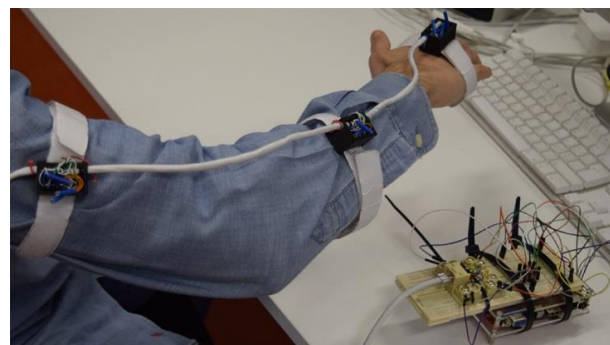


Figure 1: The Hardware Prototype

3.3.2. Independent Variable

The variable that is studied is the Sample Rate of the tracking of the arm movements. The

magnetometer is set, permanently, at its maximum (8 Hz), because it is a value much lower than the values we were studying. The accelerometer and gyroscope Sample Rate is then set at 15, 30, 60, 90 and 120 Hz and the experiment is run.

3.3.3. Experimental Task

The user is presented with a virtual avatar, viewed from a first-person perspective (Fig. 2 and Fig. 3 illustrates this). Now the user can play around with the avatar for 30 seconds. After this period, the user's first task is started. The order of the experiments was alternated between users to avoid biasing the results based on learning effect. The Sample Rate order was always chosen by the random list generator from www.random.org/lists.

The following two tasks were crafted in a way that allow us to probe both slow (balancing a ball) and fast movements (hitting the bears in succession) and study how this affects the sample rate that the users feel they need and their measured performance at each sample rate.

3.3.3.1. Task 1 – Balance Ball

The user was requested to move the arm, so it stays on a starting position. There, a ball was dropped after a short count down. The user was then required to balance the ball for as long as he can. A perfectly vertical shadow indicated where the ball would fall. This was used as an aid to the lack of good depth perception to the user, where the user is expected to use depth on his movements. This task was chosen to evaluate conditions where most concentration lies on controlling an object on a slow and predictable environment (the ball responds only to the gravity and the forces the user applies). We measured the time the user is able to balance the ball. The test is repeated for 10 iterations and then the Sample Rate is, randomly switched. On each switch, the user answers the questionnaire.



Figure 2: The first task: Balance Ball

3.3.3.2. Task 2 – Whack-a-Bear

The user is requested to hit a teddy bear that spawns on a random position on top of a table, as if playing a game of “Whack-a-Mole”. The spawn position was set such as it could never spawn too close to the previous position so a double hit could happen. This task was chosen to evaluate conditions where fast, far away, and precise movements are required due to the semi-randomness of the environment. The bear spawns 30 times for each Sample Rate, which is then randomly switched. On each switch, the user answers the questionnaire.

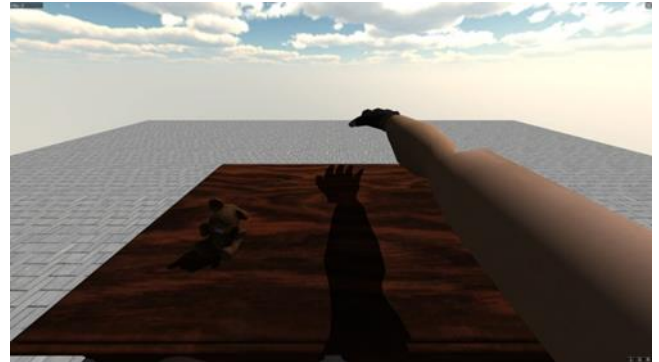


Figure 3: The second task: Whack-a-Bear

3.3.4. Data Extraction and Analysis

For the “Whack-a-Bear” experiment, we log the time it took for the user to touch the bear. For the “Balance Ball” experiment, we log for how long the user was able to balance the bear. The performance data was then analysed by calculating the average, standard deviation and also by comparing the averaged values of higher and lower Sample Rates. After those preliminary tests and analysis, we performed T-Tests of Student on each Sample Rate pair for the Balance Ball experiment time that the user was able to keep the ball in balance and for the Whack-a-Bear experiment time between bear hit. For each T-Test we also calculated the P-Value, for a two tail and the Effect Size. The used confidence interval was 95%.

The subjective data was averaged, and its standard deviation was calculated. We then proceeded to analyse discrepancies in the data (for instance, if the Sample Rate increases but the users think it was slower or harder, this indicates that the user may not be able to accurately notice (or, at least, express) what changed. We then averaged all the increases and decreases of the Sample Rates and performed the previous stated analysis.

4. RESULTS

4.1. Balance Ball

The tests indicate a good effect when going from 15 to 120 Hz ($t=0.07$, $P=0.95$, effect size=0.01) but

also a strong “non-effect” when going from 30 to 60 Hz ($t=2.01$, $P=0.05$, effect size=0.31). All the remaining data show no statistical relevance. When comparing the relation between the time it takes to complete the tasks on the different sample rates, we see no tendency in the data.

4.2. Whack-a-Bear

The tests show only a strong “non-effect” when going from 15 to 30 Hz ($t=2.01$, $P=0.0$). There seems to be no tendency in the data when comparing the times.

4.3. Questionnaire

For the Whack-a-Bear experiment, the user can always notice a good improvement (2.0) when going up from 15 Hz. But they can also notice a slight improvement (~0.5) when going down from 120 Hz, even to the other extreme of 15 Hz. For the remaining of the results, the users can guess about 53% of the time. As for the Balance Ball, there is a less pronounced effect of the user noticing an improvement by going either direction from one extreme to the other. The user is also more prone to correctly “guess” (65%) which direction the Sample Rate changed to, even if just slightly. There seems to be a bias towards positiveness on all the results, regardless of the tested direction.

5. DISCUSSION

The user performance values mostly indicate that no effect is present. When they indicate otherwise, those results are denied by other results. Since it makes no sense finding an improvement in both directions, those values may be disregarded as noise. This possibility is corroborated by the fact that the tests that may indicate an effect are those with the biggest standard deviation. This may be caused by an accumulated learning effect.

When looking at the relation between two Sample Rates for the Balance Ball, we can, easily, see that the ratio is split between being what it is expected (improving when increasing the Sample Rate and degrading when lowering the Sample Rate) and what is not (the opposite). In fact, it lays on the 50% mark. The Whack-a-Bear is more as expected marking at 20% contradictory results. Still, if we average all the relations, in both experiments we end up with a value of 1.0 for each which indicates that, on average, the performance is indifferent to the Sample Rate, for the tested interval. The indifference in performance was expected when the experiment was designed but only for the higher values. Having no performance difference between 15 Hz and 120 Hz came out as a surprise and one that is not easy to explain. We speculate that either the Madgwick’s algorithm is good enough to

compensate for the lower Sample Rate or that we did not design the experiment in a way that would reach a breaking point in speed and/or precision. Still, the reason could be that, in fact, the optimum value is, indeed, around or below 15 Hz.

As for what the user feels, a correct guessing of 53% and 65% for Whack-a-Bear and Balance Ball, respectively, seems to indicate that the user may not be sure what he is really feeling. It could also indicate that the user was not able to, correctly, communicate what he really felt or that the questionnaire was ill built or incomplete. There is also a bias towards positiveness that raises some red flags. This bias could be explained by either the user feeling a need to find an improvement, even if there is not one present or by the learning effect being strong enough for the user to confuse learning with technical improvement.

By adding personal remarks from the users’ interactions during the experiments, we can state that we noticed that some users were really able to correctly and consistently guess the direction of the Sample Rate change. But those users represent a small portion of the whole sample. Unfortunately, we did not take note of the exact number of users. On the other hand, the majority of the users showed no clue of whether the Sample Rate had increased or decreased. This led us to conclude that there may be characteristic or subpopulation that has higher sensitivity to a Sample Rate change. As of this moment we were not able to identify what characteristic or subpopulation it may be.

6. CONCLUSION

Performance-wise we found no evidence that changing the Body Tracking Sample Rate would change the user performance when performing tasks, be it slow or fast or even precise. We found, however, that a small group of users may notice the change in the body tracking Sample Rate. We did not find where the exact threshold is, how strong that effect is and what makes those users being able to notice the body tracking Sample Rate change. This leaves a good margin for developers to save energy and bandwidth when tracking arms.

ACKNOWLEDGMENTS

This research was partially funded by IDERAM through grants no. M1420-01-0247-FEDER-000019 and M1420-01-0247-FEDER-00003

REFERENCES

- H. Dejnabadi, B. M. Jolles, and K. Aminian. (2005) A New Approach to Accurate Measurement of Uniaxial Joint Angles Based on a Combination of Accelerometers and Gyroscopes. *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 8, pp. 1478–1484.
- H. J. Witchel, C. Westling, A. Healy, N. Chockalingam, and R. Needham. (2012) Comparing four technologies for measuring postural micromovements during monitor engagement. p. 189.
- K. Lorincz, B. Chen, G. W. Challen, A. R. Chowdhury, S. Patel, P. Bonato, and M. Welsh. (2009) Mercury: a wearable sensor network platform for high-fidelity motion analysis. p. 183.
- R. Aylward and J. A. Paradiso. (2007) A compact, high-speed, wearable sensor network for biomotion capture and interactive media. p. 380.
- S. O. H. Madgwick, A. J. L. Harrison, and R. Vaidyanathan. (2011) Estimation of IMU and MARG orientation using a gradient descent algorithm. pp. 1–7.
- T. Cloete and C. Scheffer. (2008) Benchmarking of a full-body inertial motion capture system for clinical gait analysis. pp. 4579–4582.
- T. Cloete and C. Scheffer. (2010) Repeatability of an off-the-shelf, full body inertial motion capture system during clinical gait analysis. pp. 5125–5128.
- X. Wei, P. Zhang, and J. Chai. (2012) Accurate realtime full-body motion capture using a single depth camera. *ACM Transactions on Graphics*, vol. 31, no. 6, p. 1.