A STUDY OF THE NOISE-LEVEL IN TWO INFRARED MARKER-BASED MOTION CAPTURE SYSTEMS

Alexander Refsum Jensenius,[♮] Kristian Nymoen,[♯] Ståle A. Skogstad,[♯] Arve Voldsund[♮♯]
[♮]University of Oslo, Department of Musicology
[♯]University of Oslo, Department of Informatics
`{alexanje, arvevo}@imv.uio.no, {krisny, savskogs}@ifi.uio.no`

ABSTRACT

With musical applications in mind, this paper reports on the level of noise observed in two commercial infrared marker-based motion capture systems: one high-end (Qualisys) and one affordable (OptiTrack). We have tested how various features (calibration volume, marker size, sampling frequency, etc.) influence the noise level of markers lying still, and fixed to subjects standing still. The conclusion is that the motion observed in humans standing still is usually considerably higher than the noise level of the systems. Dependent on the system and its calibration, however, the signal-to-noise-ratio may in some cases be problematic.

1. INTRODUCTION

In our research on various types of *music-related movement*, including the movements of both performers and perceivers, we are now moving towards exploring *micromovements*. Micromovement is here defined as displacements smaller than a centimetre, and can be anything from the smallest movements produced by musicians to the noise-like patterns observed when people try to stand still [1, 2].

Fortunately, exploring micromovements is nowadays possible using various types of motion capture systems. Both camera based and sensor based motion capture systems offer theoretical sub-millimetre precision and accuracy at high sampling rates [3], albeit with some trade-offs in either flexibility or spatial drift [4]. The question, though, is whether this type of data quality can be achieved in practice, both inside and outside of a lab environment.

In a previous study we have compared one inertial (Xsens) and one infrared (Qualisys) motion capture system, and discussed some of their benefits and limitations when it comes to capturing human motion in musical contexts [5]. In another study we have compared the quality of the motion sensor data obtained from a mobile device (iPod Touch) with the corresponding data from an infrared motion capture system (Qualisys) [6]. Both of these studies have focused on the quality of the data observed when moving with normal-sized motion, and the latency and drift observed when streaming the data to a remote computer. We have found that the spatial accuracy/precision and temporal speed/latency of all these systems are satisfactory for many musical applications, e.g. our *SoundSaber* [7]. Now, however, as we are working on the scientific and artistic exploration of music-related micromovements, a high spatial resolution becomes more important.

The fastest and most accurate motion capture system we have access to is an infrared marked based system (IRMoCap) from Qualisys. In this paper we look at the noise-level that can be found in this system, as compared to human micromovement, and which factors that influence the noise-level. We are also interested in seeing how the data from an affordable IRMoCap system (OptiTrack) compares to the Qualisys data. We do not consider the systems from Qualisys and OptiTrack to be competing products, but rather as being on each side of a continuum of cost and size. This also reflects how we use the systems. In the lab we work with the Qualisys system, since it provides the highest data quality, sampling rate and stability. For activities outside the lab, however, we typically use the OptiTrack system, since it is more portable and takes up less (visible) space in the environment in which it is set up.

The paper starts with an overview of the experimental methods used. Then results from the different data sets are presented, followed by a discussion of various features that may influence the noise level of IRMoCap systems.

2. EXPERIMENTAL SETUP AND METHOD

Data for this paper were recorded in three different laboratories at the University of Oslo during 2011 and 2012. Two different IRMoCap systems were used:

- Qualisys: 9 Oqus 300 cameras [1]
- Naturalpoint: 8 OptiTrack FLEX:V100 cameras [2]

Recordings were done using the two systems' native software (Qualisys Track Manager 2.7 and NaturalPoint Arena 1.7). No post-processing was applied to the data besides labelling of the markers. Motion capture data files were exported in TSV format (Qualisys) and C3D (OptiTrack), and these files were imported and analysed in Matlab using the MoCap Toolbox [8].

---

[1] `http://www.qualisys.com`
[2] `http://www.naturalpoint.com`

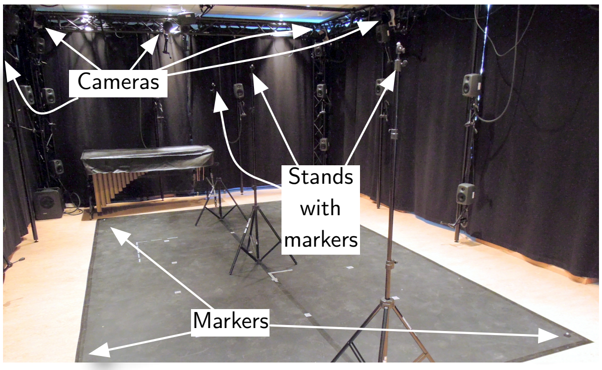

Figure 1. Picture from the lab used to record Data set #2 and most of #3, showing one of the setups for the study.

2.1 Data sets

The analyses in this paper are based on data from three separate data sets:

- Data set #1: 6 recordings of one marker lying still on the floor, and markers on the neck and right foot of 2 subjects standing still [1]. Recorded with the Qualisys system in a setup covering a space of 5 m x 7 m x 3 m, and with a sampling rate of 100 Hz and duration of 10 minutes.
- Data set #2: 25 recordings of one marker fixed to a pole standing still, and markers on the head of 3 subjects standing still [2]. Recorded with the Qualisys system in a setup covering a space of 6 m x 8 m x 3.5 m (Figure 1), and with a sampling rate of 20 Hz and duration of 10 minutes.
- Data set #3: 50 recordings of multiple markers lying still on the floor or fixed to poles standing still. Recorded with both the Qualisys and OptiTrack systems, in different rooms, with different camera setups, and with different sampling rates and durations.

To give an impression of how the data sets look like, Figure 2 shows plots of XYZ positions (centred around the mean value) of a 10-minute long recording of a marker fixed to a pole at head's height and a marker on the head of a person standing still (from Data set #2). The "motion" of the static marker is quite normally distributed, albeit with some drift, and can be seen as the noise in the system. In this example the noise level is clearly smaller than the micromovements of the subject's head. The question, however, is whether the noise level is always this much lower than the smallest observable human motion?

2.2 Quantity of motion

In our first study on standstill [1], we looked at the total cumulative distance travelled by a marker (in metres), i.e. the sum of all the differences between individual samples, to get a measure of how much the marker moved. While this measure works when comparing recordings of equal duration, it cannot be easily used to compare recordings with different duration. Then it is better to divide the cumulative distance by the time taken, which gives the average speed that the marker travelled with. We will use this measure to describe the *quantity of motion* (QoM) of the marker:

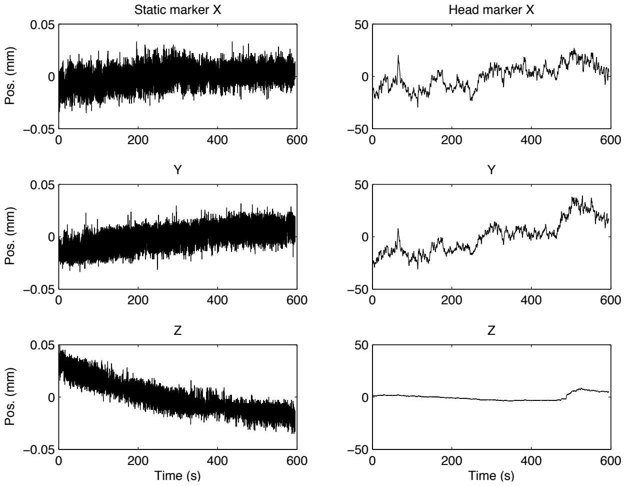

Figure 2. Example of a recording of XYZ positions (centred around the mean value) of one static marker (left) and one head marker (right), from Data set #2. Duration = 10 min, SR = 20 Hz. The scaling of the Y-axes are different for the left and right plots (by a factor of 1000).

$$QoM = \frac{\sum_{n=2}^{N} ||\mathbf{p}(n) - \mathbf{p}(n-1)||}{T} \tag{1}$$

where $\mathbf{p}$ is a 3D position vector, N is the total number of samples and T is the total duration of the recording.

One problem with using QoM as a measure in comparative studies is that the sampling rate used may influence the calculated result. A higher sampling rate will yield a higher value for the QoM, at least if there is motion happening within the frequency band that is cut away. For that reason it may be relevant to calculate a measure that is independent of time and sampling rate.

2.3 Spatial range

If we assume that the motion observed in a still marker is distributed normally (flicker noise), it may be better to look at the *size* of the volume covered by the marker, than how *much* the marker moved. One way of measuring this could be to calculate the *volume* covered, but in our experience the volume is not very practical to work with, due to the cubic nature of the function. Then we find that calculating the *spatial range* (SpR) as the magnitude of the range of each of the three dimensions is more practical:

$$SpR = \sqrt{(X_{range})^2 + (Y_{range})^2 + (Z_{range})^2} \tag{2}$$

where X, Y, Z represent the individual components of the 3D position, so that $X_{range} = X_{max} - X_{min}$, etc.

It may be argued that calculating the range is unfortunate, since for example one single spike in one dimension would dramatically increase the total measure. As such, the range represents the extremes of the distribution, not necessarily the main part of the distribution. Calculating the magnitude of the *standard deviation* of the position of a marker may then be a "better" solution, since it will effectively leave out spikes and other types of outliers in the data set. This is also the approach taken by many motion capture software packages when they present the user with a measure of the quality of the calibration. For our current study, however, we are interested in getting the full picture, and will therefore present results for the SpR and the QoM in the rest of this paper.

3. RESULTS

3.1 Noise level in Data set #1

Table 1 shows that the mean QoM of a marker placed on the neck of the two subjects in Data set #1 was 6.6 mm/s for all recordings, while the mean QoM of a marker placed on the right foot of each person was 2.2 mm/s. Correspondingly, a marker lying on the floor next to the feet "moved" 1.8 mm/s, on average. As such, the still marker moved less (on average) than the smallest observed human micromovement, but not by much. In fact, in one recording the QoM value of the foot of a person was smaller than for a marker lying still on the floor next to the foot.

To test the influence of the sampling rate on QoM, we downsampled Data set #1 from 100 Hz to 20 Hz (using the *mcresample* function in the MoCap Toolbox) before calculating the QoM. The resultant QoM values can be seen in parentheses in Table 1. The resultant mean QoM values are 5.9, 0.5 and 0.4 mm/s for the neck, foot, and static markers, respectively. These values indicate that most of the activity above 10 Hz (half the sampling rate) is most likely noise, since the QoM of the neck marker was not influenced particularly much by the downsampling process.

Table 1 also shows the mean SpR values for Data set #1. We can see that the SpR discriminates "better" between the static marker and foot marker than the QoM, and may therefore be a good measure to accompany the QoM for this type of data.

3.2 Noise level in Data set #2

Data set #1 was recorded in our old lab, and before we learned about all the different settings of the Qualisys system and the importance of the calibration. For that reason we were curious to see whether we would find any major differences in Data set #2, recorded in our new and (slightly) larger lab a year later. These recordings were done with the same Qualisys system but now mounted in a truss rig as opposed to standing on tripods.

Data set #2 was primarily recorded for studying human standstill as a case of an "inverted pendulum" [9], and we therefore recorded the micromovements of a marker placed on the head of three subjects [2]. To compare the recordings of the head markers to that of a static marker, we placed the static marker at head's height, on top of a tripod standing right next to the three subjects. This time we also decided to record with a lower sampling rate (20 Hz) to see whether/how this would influence the recordings.

Table 1. Summary of Data set #1: QoM and SpR for neck and foot markers on 2 subjects and one static marker. The table represents mean values of 6 recordings, duration = 10 min, SR = 100 Hz, no filtering applied. QoM values calculated after downsampling to 20 Hz in parentheses.

	QoM (mm/s)	SpR (mm)
Neck	6.6 (5.9)	79.6
Foot	2.2 (0.5)	3.2
Static	1.8 (0.4)	0.3

Table 2. Summary of Data set #2: QoM and SpR values for head markers on 3 subjects and one static marker. The table represents the mean values of 25 recordings, duration = 10 min, SR = 20 Hz, no filtering applied.

	QoM (mm/s)	SpR (mm)
Head	6.8	80.8
Static	0.5	0.3

A summary of the mean QoM and SpR of Data set #2 can be seen in Table 2. These results confirm our findings from Data set #1, i.e. that the micromovements of a marker placed on the upper part of a human person standing still, is clearly above the noise level observed in a static marker placed in close vicinity of the markers on the subjects. Even though Data sets #1 and #2 are not directly comparable, due to the slightly different setup and different sampling rate, it is interesting to see that both the (downsampled) QoM and the SpR values for the static markers are quite similar. As expected, the head moves a little more and has a slightly larger SpR than the neck.

The noise level found in Data set #1 and #2 is satisfactory when compared to the micromovements observed in the upper part of the body of a few people standing still in the middle of the room. However, we are currently interested in carrying out studies of larger groups of people standing still while listening to music. Ideally, we would imagine recording as many as 10–20 subjects at a time, which would require using the entire lab space for the recordings. The question, then, is whether we can expect a noise level similar to what we found in Data set #1 and #2 across the entire lab space?

4. FACTORS INFLUENCING NOISE LEVEL

To test how different factors influence the noise level, we carried out a number of recordings that will be discussed in this section. These recordings are all from Data set #3. Sections 4.1–4.5 are based on recordings with the Qualisys system, while Sections 4.6 and 4.7 are based on data data recorded with both the Qualisys and the Optitrack systems.

4.1 Calibration

When working with an IRMoCap system there are two different *volumes* that are important for the final result. As sketched in Figure 3, the *covered volume* is the part of the room that is covered by at least two of the cameras, and where a marker will be tracked in three dimensions. The

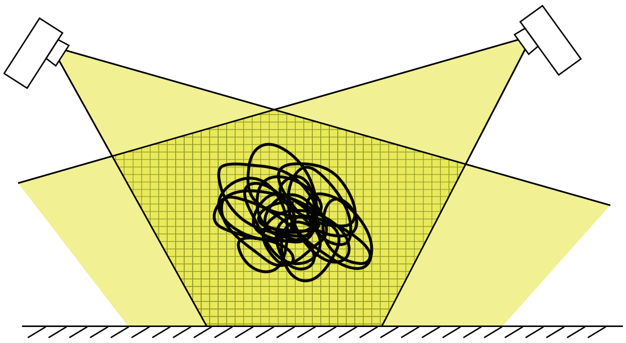

Figure 3. Illustration of how the *calibrated* motion capture volume only fills up parts of the *covered* volume.

calibrated volume, on the other hand, is the part of the covered volume that has been calibrated with a marker wand at the beginning of the recording session. If calibrated thoroughly, the calibrated volume may be almost the same size as the covered volume. But even when we try our best, we typically end up with a calibrated volume that is smaller than the covered volume.

It is also common to have some “holes” in the calibration, i.e. that a part of the calibrated volume has not been calibrated properly. This typically happens when forgetting to move the wand systematically over all parts of the volume during the calibration. The question, however, is to what extent being inside, outside or in a hole changes the noise level of a recorded marker?

To test the noise level at different locations in the space, we made two 2-minute recordings of 2 markers inside the calibrated volume, 1 marker in an uncalibrated hole, and 3 markers outside the calibrated volume. The system was calibrated to approximately the same calibrated volume for each recording, to remove any possible problems due to the calibration. The analysis of these recordings, summarised in Table 3, reveals that the mean QoM and SpR of the markers lying outside the calibrated volume are at a factor of 7–9 compared to the markers lying still inside the calibrated volume. So this indicates that it is important to keep within the calibrated volume, at least when studying micromovements.

We have been aware of the possible differences in noise level between markers inside and outside the calibrated volume, but we have not previously thought much about the holes inside the calibrated volume. As Table 3 shows, the QoM and SpR values for the marker in a hole are slightly higher than those inside the calibrated volume, but lower than those outside the calibrated volume. The difference is not large, but it should still be an indication that it is important to calibrate thoroughly, and to avoid placing markers in obvious holes.

Looking more closely at the data, we can see that the noise of the markers inside the calibrated volume is not only smaller, but also more uniformly distributed than for the markers outside the calibrated volume. Furthermore, there are almost no spikes in the data recorded within the calibrated volume, while there are several large spikes outside. It should be noted here that we are talking about spikes of approximately 1 mm, so these are still small numbers, but they are large enough to approach (and surpass) the smallest micromovements observed in humans, and should therefore be taken into account when designing experiments for studying such small movements.

Table 3. Summary of 2 recordings from Data set #3: 2 still markers inside the calibrated volume, 3 markers outside the calibrated volume, and 1 marker in an uncalibrated hole. Duration = 2 min, SR = 100 Hz.

	QoM (mm/s)	SpR (mm)
inside	1.5	0.1
hole	2.7	0.4
outside	10.5	0.9

Table 4. Values for markers inside and outside the calibrated volume, when only half of the covered volume was calibrated. Duration = 2 min, SR = 100 Hz.

	QoM (mm/s)	SpR (mm)
inside	2.1	0.1
outside	1.5	0.1

Does the size of the calibrated volume influence the noise level of markers inside the calibrated volume? It may be argued that it is better to calibrate a large volume even though recordings will only be carried out in the middle of the room, because the system will have more points distributed over a larger volume to calculate the location of the cameras. However, we have not been able to find that the size of the calibration volume significantly influences the noise level of markers inside the calibrated volume.

Out of curiosity we also tested calibrating only one half of the covered volume, splitting the room in two at the centre. Two stands with markers were placed at an equal distance from the centre, one placed one metre inside the calibrated part, and the other placed one metre inside the uncalibrated part of the room. Table 4 shows that the SpR values were equal for the two markers, and that the QoM values were actually better for markers outside the calibrated volume. While such a calibration is unlikely to be performed for normal recordings, it indicates that there are a number of factors influencing the calibration quality.

4.2 Marker size

To what extent does the marker size influence the recordings? Obviously, if the markers are too small, they will not be seen by the system at all. In our current setup we find that spherical markers of size ⌀16 mm and ⌀12 mm work well, and we have not been able to find any systematic difference in the noise level due to the marker size. Originally we thought that it would be favourable to use the same marker size for the calibration as for the recording. However, our results show that recording large markers after calibrating with a smaller wand kit, and vice versa, does not give a noise level that is statistically much different.

We have also tried recording with smaller, spherical markers (⌀7 mm) and flat markers (⌀4 mm), but these were too small to be seen satisfactorily by the system, hence leading

Table 5. Mean values of 2 markers inside the calibrated volume, recorded with different sampling rates (SR). The QoM values have been calculated after downsampling to 20 Hz, to allow for comparison. Duration = 2 min.

SR (Hz)	QoM (mm/s)	SpR (mm)
500	0.43	0.11
200	0.30	0.08
100	0.32	0.11
50	0.29	0.07
20	0.33	0.10

to dropouts in the recorded data. Adjusting the exposure threshold of the system may improve the results.

4.3 Sampling rate

Does the sampling rate (SR) influence the results obtained? To test this we carried out a series of 2-minute long recordings of 2 markers inside the calibrated the volume and with different sampling rates: 20 Hz, 50 Hz, 100 Hz, 200 Hz, 500 Hz, with 500 Hz being the maximum sampling rate of the Qualisys system. As Table 5 shows, the noise level is fairly consistent for all recordings: the mean SpR values are all ∼0.1 mm/s, and the QoM values (downsampled to 20 Hz for comparison) are also fairly consistent.

4.4 Influence of lights

IRMoCap systems work with light in the infrared spectrum, and should therefore not be influenced by regular lighting. This we have confirmed by looking at the differences in recordings done in rooms with different light types: fluorescent, halogen and LED. We have also done recordings with the lights turned off, and turning the light on/off during recordings. None of these changes in lights have influenced the recordings. What we have not tested, and that would be interesting to check at a later stage, is whether stage lights may influence the systems (if they have frequencies in the infrared spectrum).

4.5 Occlusion

As opposed to inertial systems, IRMoCap systems are heavily influenced by occlusion of markers from one or more cameras [10]. The calculation of a marker's position in three dimensions is based on triangulating the 2D positions of the marker captured with two or more cameras. Thus changes in which cameras see the marker will necessarily introduce some noise when a marker "moves" from one camera to another. But how does occlusion influence markers lying still?

To test the influence of occlusion, we recorded a set of markers lying still on the floor and fixed to poles in the space, and let one person walk around in the space during the recording. The person did not specifically try to cover up the markers, but he moved around so that some markers were occluded from some of the cameras at all points in time. For a 2-minute recording (SR = 100 Hz) this resulted in a mean QoM of 3.4 mm/s, as opposed to QoM = 1.5 mm/s for a recording with the same marker setup without any occlusion. So occlusion clearly influences the recordings, even for markers lying still.

Table 6. Results of recordings of the same marker setup with both the Qualisys and OptiTrack systems. Duration = 30 min, SR = 100 Hz.

	OptiTrack		Qualisys	
	QoM (mm/s)	SpR (mm)	QoM (mm/s)	SpR (mm)
Floor	12.5	3.8	3.2	0.68
Stand	7.8	3.9	3.6	0.57

4.6 Qualysis versus OptiTrack

Although all of the above mentioned tests have been carried out with the Qualisys system, the key findings are probably relevant for other IRMoCap systems. However, the noise levels of different systems are not so easy to compare without actually testing two (or more) systems under the same conditions. Since we are interested in using an OptiTrack system for musical activities outside the lab, it is relevant to know what type of noise level we can expect from this system as compared to the Qualisys. The specifications of the two systems reveal that there is a quite large theoretical difference in the spatial resolution of the two systems. But how do they compare in a real lab setting?

To compare the noise level of the two systems we set up the OptiTrack system in the lab in which the Qualisys system is mounted. The OptiTrack cameras were placed as close as possible to the Qualisys cameras, we used the same markers and marker location (inside the calibrated volume) for the recordings, and the same sampling rate (100 Hz). As summarised in Table 6, the OptiTrack QoM value for a marker on a stand in the centre of the calibrated volume was twice that of the Qualisys system, while a marker on the floor (also inside the calibrated volume) was more than three times larger. The SpR values reveal that the noise level of the OptiTrack is 5–6 times that of the Qualisys system. So there is a clear difference in noise level between the two systems, but, taking size and cost into account, the OptiTrack system performs very well.

4.7 Drift

As we have documented previously [10], there is a considerable drift in many inertial motion capture systems. Similar types of drift are not seen in IRMoCap systems. However, as can be seen in Figure 4, there is a slight drift over time both in the Qualisys and OptiTrack systems for a 30-minute long recording. There may be several reasons for this, e.g. the filtering happening in the system itself, that the camera sensors get warmer over time or that the cameras themselves actually move slightly over time. Still, the level of drift observed is so small that it in practice could be considered negligible.

5. CONCLUSIONS

Based on our systematic study of still markers in two IR-MoCap systems, we can conclude the following:

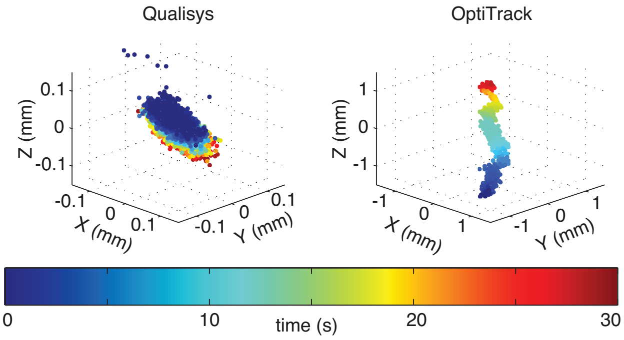

Figure 4. A slight drift can be seen in the XYZ plots of markers on a stand in the middle of the calibrated volume, measured with both the OptiTrack and Qualisys systems. Duration = 30 min, SR = 100 Hz. Notice the scale difference on the axes.

- Quantity of motion (QoM) and spatial range (SpR) can be used to describe time-dependent and time-independent noise/motion levels, respectively.
- Noise floor: the noise level observed for markers within the calibrated volume is slightly lower (OptiTrack) and considerably lower (Qualisys) than the micromovements observed in humans standing still.
- Calibrated vs. covered volume: the difference in noise level inside and outside the covered volume is fairly large. Care should also be taken to avoid uncalibrated "holes" in the calibrated volume.
- Marker size: as long as the markers are seen by the system throughout the recording, there is no clear difference between using larger or smaller markers.
- Sampling rate: the sampling rate does not seem to influence the noise level.
- Drift: there is very little drift, even for long recordings (30 minutes).
- Lights: we have not been able to document any influence of fluorescent, halogen and LED lights. Stage lights remain to be tested.
- Qualisys versus OptiTrack: The Qualisys system provides a lower spatial noise level, and at higher sampling rates, than the OptiTrack system. However, for many (musical) applications, the noise level of the OptiTrack system is low, especially if comparing to other types of motion capture solutions (inertial, regular computer vision, etc.).

The knowledge gained from this study will serve as the baseline for developing better filters and analysis methods for motion capture data, as well as for developing future studies of music-related micromovements.

Acknowledgments

This research has been funded by the Norwegian Research Council through the project *Sensing Music-related Actions* (#183180) and by the EU FP7 project *EPICS* (#257906).

6. REFERENCES

[1] A. R. Jensenius and K. A. V. Bjerkestrand, "Exploring micromovements with motion capture and sonification," in *Proceedings of the Second International ICST Conference on Arts and Technology (ArtsIT)*, Esbjerg, Denmark, 2011.

[2] A. R. Jensenius, K. A. V. Bjerkestrand, and V. Johnson, "How still is still? exploring human standstill for artistic applications," *International Journal of Arts and Technology*, 2012 (in review).

[3] G. Bishop, G. Welch, and B. Allen, "Tracking: Beyond 15 minutes of thought," *SIGGRAPH Course Pack*, 2001.

[4] T. Cloete and C. Scheffer, "Benchmarking of a full-body inertial motion capture system for clinical gait analysis," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, aug. 2008, pp. 4579 –4582.

[5] S. A. v. D. Skogstad, K. Nymoen, and M. E. Høvin, "Comparing inertial and optical mocap technologies for synthesis control," in *Proceedings of Sound and Music Computing*, Padova, Italy, 2011, pp. 421–426. [Online]. Available: http://smcnetwork.org/system/files/smc2011_submission_124.pdf

[6] K. Nymoen, A. Voldsund, S. A. v. D. Skogstad, A. R. Jensenius, and J. Tørresen, "Comparing motion data from an iPod touch to a high-end optical infrared marker-based motion capture system," in *Proceedings of the International Conference on New Interfaces For Musical Expression*, Ann Arbor, MI, 2012.

[7] K. Nymoen, S. A. Skogstad, and A. R. Jensenius, "Soundsaber - a motion capture instrument," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Oslo, Norway, 2011, pp. 312–315. [Online]. Available: http://urn.nb.no/URN:NBN:no-29584

[8] P. Toiviainen and B. Burger, *MoCap Toolbox Manual*. University of Jyväskylä, 2010. [Online]. Available: http://www.jyu.fi/music/coe/materials/mocaptoolbox/

[9] I. Loram, H. Gollee, M. Lakie, and P. Gawthrop, "Human control of an inverted pendulum: Is continuous control necessary? Is intermittent control effective? Is intermittent control physiological?" *The Journal of Physiology*, vol. 589, no. 2, pp. 307–324, 2011.

[10] S. A. Skogstad, K. Nymoen, Y. D. Quay, and A. R. Jensenius, "OSC implementation and evaluation of the Xsens MVN suit," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Oslo, Norway, 2011, pp. 300–303. [Online]. Available: http://www.nime2011.org/proceedings/papers/G23-Skogstad.pdf