

Comparative Analysis of OptiTrack Motion Capture Systems

Joshua S. Furtado, Hugh H. T. Liu

University of Toronto Institute for Aerospace Studies
University of Toronto
Toronto, Canada
joshua.furtado@mail.utoronto.ca

Gilbert Lai, Herve Lacheray, Jason Desouza-Coelho

Quanser
Markham, Canada
gilbert.lai@quanser.com

Abstract— A comparative analysis of four different OptiTrack motion capture systems is provided in this paper. Tracking accuracy, workspace volume, marker size and camera range are the main criteria used for comparison. The tracking accuracy is tested using Quanser's linear motion platform by comparing the measured position of the cart with the position obtained using the encoder. The workspace volume limits are obtained by flying a LiteHawk Neon drone in circles of increasing radius at different heights until tracking fails. The relationship between marker size and camera range is obtained through theoretical calculations. Experimental as well as theoretical results are presented, illustrating the performance of these four systems. These will serve as a baseline to select the right motion capture system for any particular application.

Keywords- *Motion capture; OptiTrack; Optical-Passive*

I. INTRODUCTION

Motion capture (mocap) is the process of recording the movement of objects or people. It involves measuring the position as well as orientation of the objects or people in physical space [1]. The technology was originally developed for gait analysis in the life science market but is now used in a wide variety of other fields. Some of these include computer graphics animation, robotics and military use.

Over the past few decades, a lot of advancements have been made in motion capture technology resulting in many different approaches such as:

- **Optical-passive:** Retroreflective markers are tracked by a set of infrared cameras. It is the most practical and commonly used method in the industry [2]. The motion capture systems selected for comparative analysis in this paper are all optical-passive.
- **Optical-active:** Infrared LED markers are tracked by a set of infrared cameras. Note that a suitable power supply source is required to power active markers.
- **Video:** Software is used to track the object motion in each frame. Depending on the object detection and tracking algorithm used, the accuracy of the results

may vary. This technique is less accurate than the marker-based solutions.

- **Inertial:** Independent of cameras, this technique uses inertial measurement sensors on the subject. Short term accuracy is a major drawback.

Optical-passive motion capture systems are widely used for validation and control of computer vision and robotics in an indoor laboratory setting. They have great practical value in indoor mobile robotics applications. Hence, choosing the right motion capture system for a particular application is extremely important. In this paper, a series of systematic tests are performed using four different OptiTrack motion capture systems to characterize their performance. The results obtained from these tests will provide a baseline for users to choose the right motion capture system for any computer vision or robotics application based on tracking accuracy, size of workspace, marker size and camera range.

This paper is organized as follows. In Section II, the four different OptiTrack motion capture systems selected for evaluation are presented. In Section III, the experimental setup used to test the performance of these motion capture systems is described and the test procedure is explained in detail. The results obtained by performing the test using the four different systems are showcased in Section IV. Finally, the conclusions drawn from this work are presented in Section V.

II. MOTION CAPTURE SYSTEMS

OptiTrack makes high-performance motion capture systems at an affordable price. A single OptiTrack motion capture system may track the 6 degrees of freedom (DOF) pose of one or more objects in the workspace. Objects tracked in space need to be defined as rigid bodies. Each rigid body is a cluster of reflective markers arranged in a unique configuration. Fig. 1 shows two quadrotors (a LiteHawk Neon and LiteHawk Highroller) defined as rigid bodies, one using 1/2-inch by 1/2-inch square markers and the other using 5/8-inch diameter spherical markers. OptiTrack manufactures spherical markers with diameters ranging from 1/4-inch to 3/4-inch. Additionally, OptiTrack also does fabricate reflective 1/3-inch diameter dots and 1/2-inch by 1/2-inch squares. Depending on the size and limitations of the rigid body, appropriate size markers may be







Figure 1. LiteHawk Neon and Highroller quadrotors tagged with retroreflective markers

chosen. For example, we use squares on the LiteHawk Neon due its payload capacity and spherical markers on the LiteHawk Highroller because of its fuselage design.

In this paper, four different OptiTrack systems are chosen for comparative analysis, comprising: Flex 3 (x12), Flex 13 (x6), Prime 13 (x6) and Prime 17W (x6). Table 1 provides detailed specifications for the cameras used in each of these systems. The specifications provided are obtained directly from the manufacturer's website [3]. The number of cameras for each of the systems are chosen to cover roughly the same workspace area and to be able to track rigid bodies using square markers. Note that the dot and square markers expose less retroreflective surface area compared to the spherical markers. The Flex 3 cameras have a much lower resolution compared to the rest. The ethernet connectivity of the Prime series allows for a faster frame rate and lower latency compared to the Flex series. All cameras use wide angle lenses. The system cost is the number of cameras times the cost per camera. However, this does not include the costs of the hubs, proprietary software and the ground control station as they are much smaller compared to the cost of the cameras.

TABLE I. MOTION CAPTURE SYSTEMS CHOSEN FOR COMPARATIVE ANALYSIS

	OptiTrack Flex 3 (x12)	OptiTrack Flex 13 (x6)	OptiTrack Prime 13 (x6)	OptiTrack Prime 17W (x6)
				
Resolution	640 × 480 (0.3 MP)	1280 × 1024 (1.3 MP)	1280 × 1024 (1.3 MP)	1664 × 1088 (1.7 MP)
Frame rate	100 FPS	120 FPS	240 FPS	360 FPS
Lens	4.5 mm	5.5 mm	5.5 mm	6 mm
Horizontal FOV	46.2°	56°	56°	70°
Vertical FOV	34.7°	46°	46°	49°
Connectivity	USB 2.0	USB 2.0	Ethernet	Ethernet
Latency	10 ms	8.3 ms	4.2 ms	2.8 ms
Cost per camera	599 USD	999 USD	1999 USD	3499 USD
System cost	7188 USD	5994 USD	11994 USD	20994 USD

III. EXPERIMENTAL SETUP AND TEST PROCEDURE

The four different motion capture systems are compared using tracking accuracy, workspace volume, marker size and camera range. The tracking accuracy and workspace volume are obtained experimentally while the relationship between

marker size and camera range is derived geometrically based on the pinhole camera model.

A. Tracking Accuracy

Fig. 2 shows the experimental setup used to test for rigid body tracking accuracy. Quanser's linear motion platform is used where the cart and the two ends of the track are defined as rigid bodies, tracked by the motion capture system. The positions of the rigid bodies are attained using the motion capture system. The cart is driven on the track by a DC motor via a rack and pinion mechanism that ensures consistent and continuous traction [4]. It has an encoder with a resolution of 4096 counts/rev to measure its position. To start, the cart is placed at one of the ends and using the encoder feedback, a PID position controller is used to move the cart with a fixed amplitude at increasing frequency. The position of the cart on the track is obtained using its initial position, the track unit vector (defined using the positions of the two ends of the track) and the encoder count. The calculated position is then compared with the cart position obtained from the motion capture system (to test the accuracy of the four different motion capture systems). For each camera system, three sets of tests are done in three different table orientations and for each orientation, data for five experimental runs are logged.

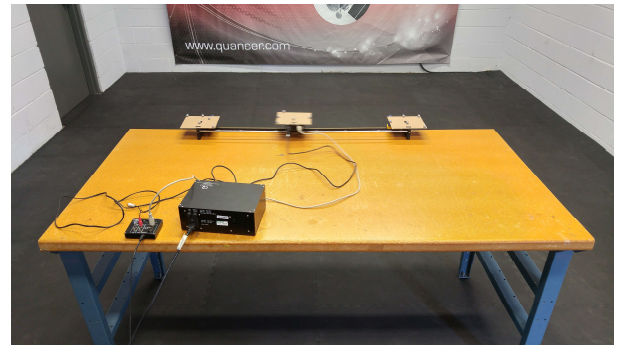


Figure 2. Experimental setup to test tracking accuracy

Let P_{CC} and P_{CE} denote the positions of the cart measured using the motion capture system and calculated based on the encoder count respectively. Let P_{E1} and P_{E2} denote the positions of the two ends of the cart obtained using the motion capture system.

$$\hat{t} = \frac{P_{E1} - P_{E2}}{\|P_{E1} - P_{E2}\|_2} \quad (1)$$

where \hat{t} is the track unit vector. The displacement of the cart from its initial position, d can be calculated using the encoder count, c . The effective linear tracking resolution, $k = 2.2749 \times 10^{-5}$ m is based on the gear ratio, pitch and encoder resolution.

$$d = kc \quad (2)$$

The position of the cart at any time, t can then be calculated as follows:

$$P_{CE}(t) = P_{CC}(t_0) + td. \quad (3)$$

The tracking error, e at every time instant, t is then given by

$$e(t) = P_{CE}(t) - P_{CC}(t). \quad (4)$$

Fig. 3 shows the experimental data for a single run using the OptiTrack Prime 13 motion capture system. The position controller moves the cart back and forth with a fixed amplitude of 0.15m at increasing frequency. The x , y and z positions of the cart measured using the motion capture system (shown in blue) are compared with those obtained using the encoder count (shown in orange). Fig. 4 shows the errors in position along the x , y and z direction for the same run.

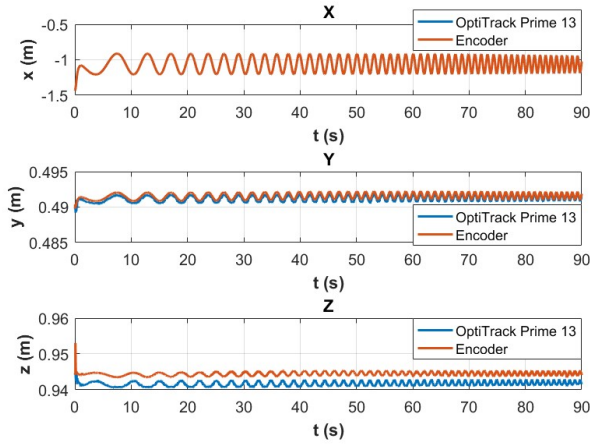


Figure 3. Measured (OptiTrack Prime 13) vs. encoder position of cart

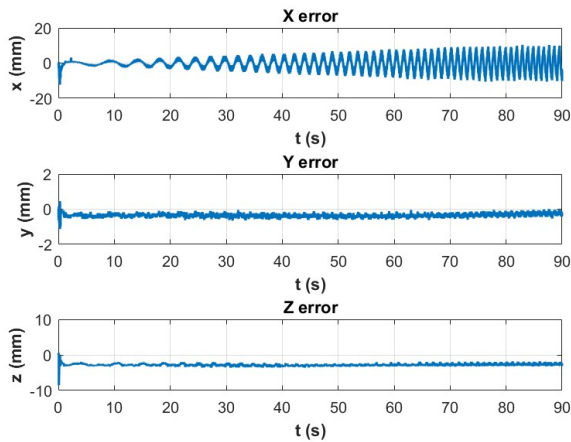


Figure 4. Tracking error (OptiTrack Prime 13)

It is important to note the tracking accuracy of the motion capture system for different cart velocities, accelerations and jerks. The velocity of the cart, V_{CC} is obtained using a fourth order forward finite difference approximation [5] of the cart position given by

$$V_{CC}(t) = \frac{\begin{pmatrix} -25P_{CC}(t) + 48P_{CC}(t+h) - 36P_{CC}(t+2h) \\ +16P_{CC}(t+3h) - 3P_{CC}(t+4h) \end{pmatrix}}{12h} \quad (5)$$

where h is the sample time. A sample time of 0.001s is used in all experimental runs. Similarly, acceleration and jerk are obtained by differentiating the approximated velocity and acceleration respectively.

B. Workspace Volume

The workspace of each motion capture system is the volume in which a rigid body can be successfully tracked. To estimate the workspace volumes of the four different motion capture systems, a LiteHawk Neon drone is flown around in the workspace to test its limits. A PID controller is used to control the position of the drone in the workspace using measurements from the motion capture system as feedback. The drone is flown around in circular trajectories of increasing radius starting at an altitude of $z = 0.25\text{m}$ until tracking is lost. This is then continued for altitudes all the way up to $z = 3\text{m}$ with increments of $\Delta z = 0.25\text{m}$. By logging the position of the drone continuously during the experiment, the workspace volume can be estimated.

C. Marker Size and Camera Range

OptiTrack manufactures spherical markers of sizes ranging from 1/4-inch to 3/4-inch in diameter. Limited by camera resolution and Field Of View (FOV), there is a maximum range to see markers of a particular size. Based on camera imaging geometry, the number of pixels seen in a marker of known size at a distance can be calculated. Knowing the marker diameter, d , the maximum cross-sectional area for each marker can be calculated using

$$A_{CS} = \frac{\pi}{4} d^2. \quad (6)$$

If $r = a \times b$ pixels is the camera resolution, the aspect ratio is

$$AR = \frac{a}{b}. \quad (7)$$

Given the horizontal FOV (h_{FOV}), vertical FOV (v_{FOV}) and focal length of the camera (f), the diagonal FOV can be calculated as follows

$$FOV = 2 \operatorname{atan2} \left(\frac{f \sqrt{\tan^2 \left(\frac{v_{FOV}}{2} \right) + \tan^2 \left(\frac{h_{FOV}}{2} \right)}}{2}, f \right). \quad (8)$$

The area seen by the camera at a distance, h is

$$S = \frac{\left(2h \tan \left(\frac{FOV}{2} \right) \right)^2}{AR + \frac{1}{AR}}. \quad (9)$$

The number of pixels seen in a marker at a distance, h is

$$N = \frac{abA_{CS}}{S} \quad (10)$$

IV. RESULTS

The experimental and theoretical results for the four different motion capture systems are presented in this section. Average error for different frequencies, velocities, accelerations and jerks of the cart are provided in the tracking accuracy subsection. Top and sectional views of the workspace volumes are shown in the workspace volume subsection. In the last subsection, the number of pixels seen in a marker by a single camera is related to its range based on the calculations in the previous section. Fig. 5 shows a scatter plot of the average error plotted against system cost. Note that the Flex 3 motion capture system is comprised of twelve cameras (twice as many as the number of cameras in the other motion capture systems). The Flex 3 cameras are mounted at a lower height to be able to track the 1/2-inch by 1/2-inch squares given their lower resolution compared to the rest of the cameras. The area seen by the cameras is thus much lower. Hence, twelve cameras are used to cover roughly the same workspace volume. The system costs reflect important characteristics of the cameras such as resolution, frame rate and latency based on the average tracking error of the four systems.

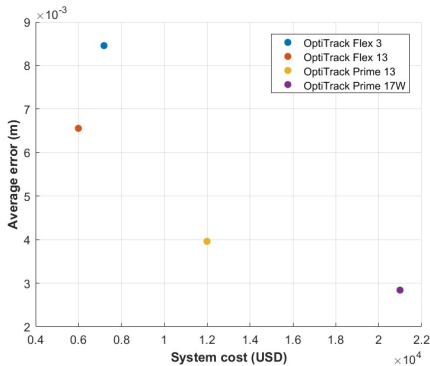


Figure 5. Average error vs. system cost

A. Tracking Accuracy

Fig. 6 shows the average error plotted against the frequency at which the cart moves back and forth along the track with an amplitude of 0.15m. The rise in average error with an increase in frequency follows a linear trend. The rate at which error increases with increase in motion frequency is a function of the frame rate and latency of the camera system. Figs. 7, 8 and 9 show plots of the average error vs. cart velocity, acceleration and jerk respectively. Note that the cart velocity, acceleration and jerk are derived using fourth order forward finite difference approximation of the cart position measured using the motion capture system. This is only an approximation and is subject to error. Straight lines were used to fit the data best from scatter plots of the average error vs. cart velocity, acceleration and jerk. The average error increases linearly with increase in cart velocity, acceleration and jerk. Once again, it can be noted that

Prime series have a lower tracking error compared to the Flex series because of higher frame rate and lower latency.

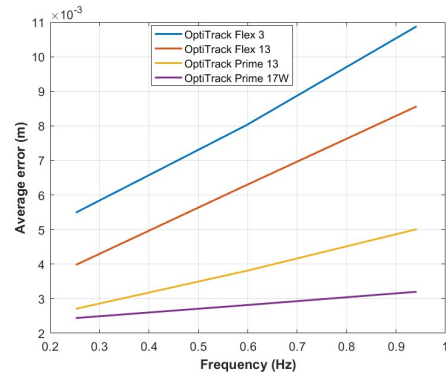


Figure 6. Average error vs. frequency

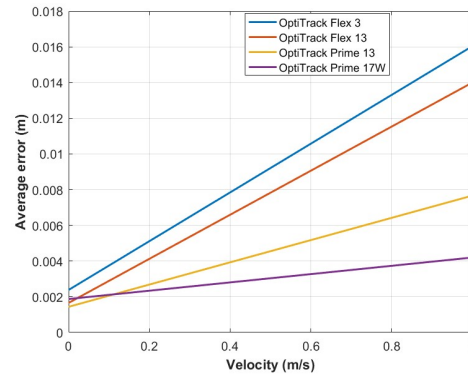


Figure 7. Average error vs. velocity

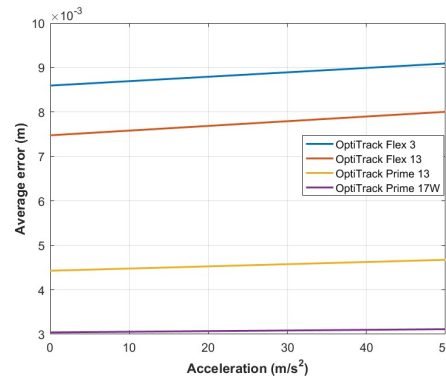


Figure 8. Average error vs. acceleration

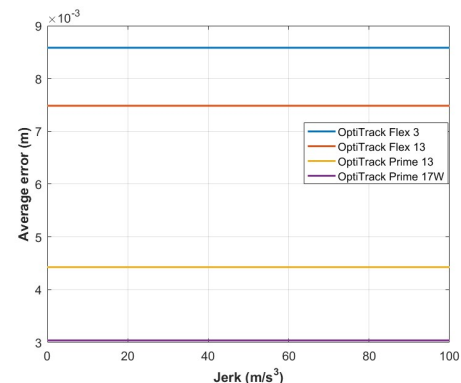


Figure 9. Average error vs. jerk

B. Workspace Volume

The limits of the workspace volume for each of the four motion capture systems were tested by flying the LiteHawk Neon drone in circular trajectories of increasing radius at different heights. Fig. 10 shows the top and sectional views for each of the four different motion capture systems. The shape of the workspaces roughly resembles that of a dome. The top views show the circles at different heights within which the drone can fly while being tracked successfully. It can be noted that the flying radius in each transverse plane decreases with increase in height. The workspace volume depends on many factors such as the positioning of the cameras, resolution, focal length and FOV.

C. Marker Size and Camera Range

Fig. 11 shows the relationship between marker sizes and range for the different camera systems. In the first plot, the number of pixels as a function of distance is seen for different marker sizes using a Flex 13 camera. The larger the marker, the more pixels are seen by the camera and hence can be seen from a longer range. The number of marker pixels seen by a camera decreases exponentially with increase in range. In the second plot, comparing the four different cameras using the same size marker at a fixed distance, the Flex 13 and Prime 13 capture the most number of pixels closely followed by the Prime 17W leaving the Flex 3 with the least. For a marker to be detected, at least 4 pixels are needed to be seen by a camera. The black line in the plots indicates this threshold.

V. CONCLUSIONS AND RECOMMENDATIONS

The Prime series cameras have a lower tracking error for aggressive rigid body manoeuvres due to higher frame rate and lower latency. High frame rate and low latency are useful in estimating rigid body velocity, acceleration and jerk more accurately. The OptiTrack Prime 17W cameras offer maximum workspace volume due to wider angle lens with bigger FOV. Using a camera with a larger FOV will result in a bigger workspace volume. The OptiTrack Flex 13, Prime13 and Prime 17W cameras offer longer marker range due to higher resolution and longer focal length. The camera resolution is important based on the type of markers used to identify rigid bodies in space. It will determine the minimum separation distance between the markers to be identified uniquely, and the minimum size of marker needed at a particular distance from the camera.

REFERENCES

- [1] Vicon Motion Systems Ltd. "What is motion capture." *VICON*, www.vicon.com/what-is-motion-capture.
- [2] Nogueira, Pedro. "Motion Capture Fundamentals: A Critical and Comparative Analysis on Real-World Applications" (2011): 1-12.
- [3] NaturalPoint, Inc. DBA OptiTrack. "Compare Cameras." *OptiTrack*, optitrack.com/hardware/compare.
- [4] Quanser. "Linear Motion Platform Product Categories." *Quanser*, www.quanser.com/products_category/linear-motion-platform/.
- [5] Fornberg, Bengt. "Generation of Finite Difference Formulas on Arbitrarily Spaced Grids." *Mathematics of Computation* 51.184 (1988): 699-706.

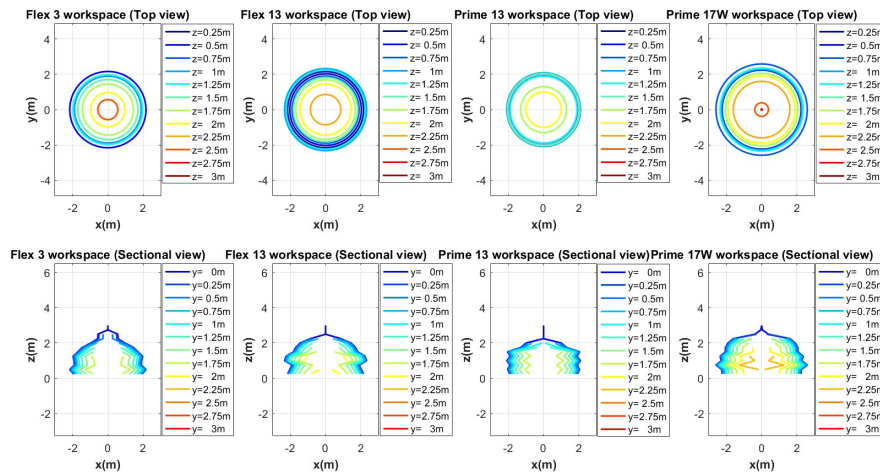


Figure 10. Workspace volume

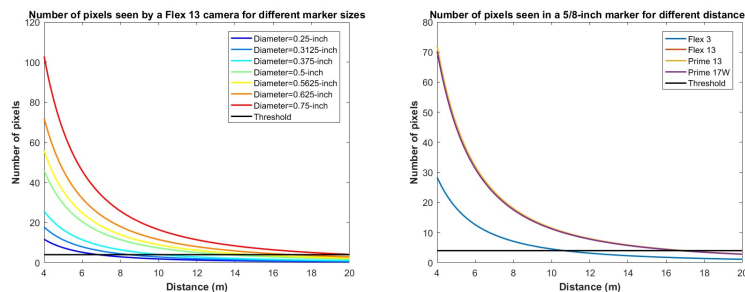


Figure 11. Number of pixels vs. distance