Affordable Low Cost Optical Tracking for VR
Technical Report

Christopher Parker
cparker@cs.uct.ac.za

Bertus Labuschagne
blabusch@cs.uct.ac.za

Russell Joffe
rjoffe@cs.uct.ac.za

ABSTRACT

A low-cost optical tracking system can lead the way to affordable realism and provide an alternative communication interface to conventional user-computer interaction methodologies. The main aim of this project was to prove that it is possible to create a low-cost optical tracking system that is easy to use, provides a high level of accuracy and has low latency

Although many optical tracking systems exist today, their high cost leaves such technology out of reach of many people. By aiming to reduce the cost of the system to as low as a couple of hundred rand, this project aims to make optical tracking technology available to a wide variety of users.

1. INTRODUCTION

User tracking is essential for a free experience of VR. It was the aim of this project to implement a system whereby a user's movements can be tracked and converted into 3D transformations that allow the user to navigate in a virtual environment in a natural way. The system operates under low-lighting conditions and was constrained by a limited budget.

UCT's VR room is based on fish tank VR. Fish tank VR is characterised by a limited physical display, in terms of size, whose location does not change [12]. Additionally, the display is rear-projected, so that the user can move around without creating shadows on the screen and decreasing immersion, by obstructing the projection system. The use of an active device to control movement in such an environment, such as a keyboard, mouse or joystick can be seen as an immersion inhibitor. A more intuitive and immersive control method is to track the user's position and orientation. Different forms of user tracking exist making use of different bands of the electromagnetic spectrum, such as audio/visual and infra-red. Each has their own advantages and disadvantages. We have designed an optical tracking system.

To summarise, this project has lead to the development of a low-cost optical tracking system for fish tank VR that takes all the above mentioned criteria into consideration and provides a more immersive experience to the user of the system. The following system requirements have been regarded as critical throughout the system design:

1. The system should be low-cost
2. The system should work in low-light conditions
3. The system should provide movement capabilities for six degrees of freedom
4. The system should be easy to use
5. The system should be efficient enough to minimize lag and delay
6. The system should have a moderate to high degree of accuracy

This report aims to show that the above requirements have been met or provides a detailed reasoning as to why certain requirements were not met.

2. BACKGROUND

2.1 Marker Extraction

A key feature of marker-based optical tracking is the use of physical markers placed on an object. These markers must be extracted from the images and transformed into discrete positions on the cameras' image planes.

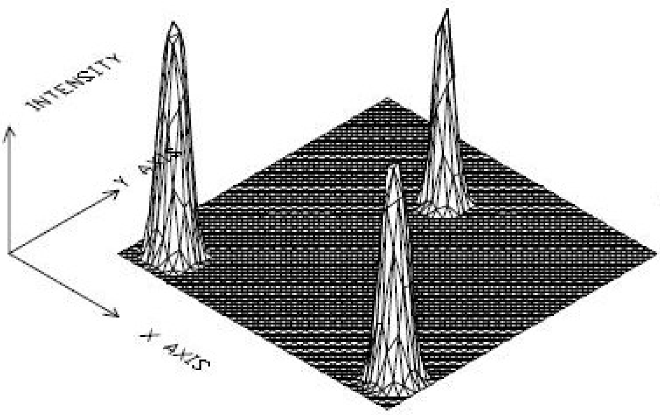

Figure 1. Ideal retro-reflective targets [**14**].

In a low-light environment, the markers, typically LED lights, will show up on the image as regions of high intensity as in Figure 1.

In "*A comparison of some techniques for the subpixel location of discrete target images*" [14], methods of determining the sub-pixel location of a target are analysed. Methods range from finding the centroid of the region or the weighted centroid to fitting the region with a normal Gaussian distribution.

Before determining the discrete location of the target, one must locate the target area or the region of interest. A simple and general way to do this is to scan the entire image captured by the camera for the target areas. If we note that a user cannot move to another location instantaneously, we can reduce the search space and increase performance. This is the approach taken in "*Search Space Reduction in Optical Tracking*" [13]. Van Liere and van

Rhijn use previously determined points of the user's motion to predict the likely 3D region in which markers will be. This 3D region is then projected onto the camera image planes and pixels within the area defined by the projection are searched first. The system is adaptive via a predictor metric, which estimates how well the system is performing.

Reducing the search space allows more time for accurate point extraction from the target area and is, therefore, highly desirable. However, if the target is not found within the search space, the entire image must be scanned and this creates variation in performance.

Mulder et al. [15] use a one-dimensional convolution filter to detect light edges (dark-to-light or light-to-dark transitions). The distances between successive light-to-dark and dark-to-light transitions are measured and those that fall within range of the expected marker size are recorded. If the number of neighbouring scanlines also falls within range of the expected marker size, then this region is marked as a region of interest. A flood fill is performed on these regions of interest to determine which pixels are members of the blob. A number of methods are employed to ignore candidate blobs which cannot be markers, such as an ellipsoid check. The ellipsoid check filters blobs based on surface area, width, height and a check that the centre of mass is central in the bounding box of the blob.

2.2 Motion Prediction

Motion Prediction has a primary purpose for reducing the latency in an optical tracking system, by allowing the system to estimate the location and orientation of the user at a given time in the future. This will then reduce the time it takes to track the user's motion. This, in turn, should reduce the latency experienced by the user.

There exist various methods for motion prediction in virtual environments. Some of these methods are adapted to work with a specific input device, i.e. either a mouse or head-tracking.

Variables which need to be considered during the prediction of motion, includes the velocity of the user motion, the characteristics of the individual user and also the user input interface. Here low velocity motion can be interpreted as slight changes in the position/rotation of the user. High velocity motion can then be seen as relatively quick changes in position/ rotation of the user.

Since each user has a unique way of moving around in an environment, it would also be prudent to take into account the characteristics of the user. These characteristics will then allow the optimization of the method for individual use.

The algorithm used depends on the input used by the virtual environment. A mouse/keyboard input only operates with 2D vectors, whereas head-motion tracking uses 3D vectors. This implies that algorithms can be adapted to be more efficient when used in conjunction with certain types of input devices.

2.2.1 The Kalman filter

A Kalman filter is a linear estimator, optimally minimizing the expected mean-square error in the state variables, i.e. location, velocity and acceleration. This filter uses the knowledge of how the environment variables have changed over time without external input. White noise processes are used to characterize the inaccuracies in both the model containing this information and the measurements made.

Even though realistically these conditions are not met in practice, the Kalman filter is still the most commonly used for two reasons. One it still performs well regardless of the fact that these conditions are violated. And two it can be described efficiently by a recursive formula. Since the filter needs to operate in real time, efficiency is of the utmost importance.

Concepts, surrounding the Kalman filter, are introduced in Welch and Bishop's "An Introduction to the Kalman Filter". Here they discuss the Kalman filter by looking into the equations of the filter.

2.2.2 Search Space Reduction

Optical Tracking requires finding locations of markers in a stream of frames from the cameras in use. This can be a very slow process and lead to an increase in latency if the search for these markers is inefficient.

In order to overcome this one can use a strategy known as search space reduction. Search space reduction allows searching for the locations in a smaller window size than the size of the frame. An exhaustive search of the frame is then only required if the marker is not in the reduced search space.

In van Liere et al [13], they discuss two different strategies for determining the search space window. The one is using a window of constant dimension and shape and the other strategy employs a dynamically shaped and sized window. They show that the latter of these performs better than using constant shaped and sized search windows.

They provide a detailed description for implementing dynamic search windows. They use a predictor to predict where the marker might be located and then create a window around those estimated points. There are two ways for achieving this. Firstly it is possible to do the prediction in the constructed 3D model, and then project these onto the 2D camera viewing planes. The other option is to do the predictive analysis in the 2D planes. However the first strategy is better since it is then possible to also take into account the angular velocity of the markers being tracked.

2.3 Presence

Virtual environments are primarily built to invoke feelings of presence. In order to quantify this presence, it is necessary to use sound presence measurement techniques. Such techniques can be classified according to the data they produce thereby concluding whether their use would lead to a good measure of presence capable of proving our hypotheses. In order to do this, it is widely accepted to group the presence measurement techniques into two broad categories, namely subjective and objective techniques. These two categories can further be subdivided into qualitative and quantitative techniques. The various techniques that fall into these categories will be classified in the following sections.

2.3.1 Subjective presence measurement

The Collins English dictionary describes the word subjective as follows:

> *of, relating to, or emanating from a person's emotions prejudices etc. subjective views.*

Existing subjective presence measurement techniques aim to gauge the effectiveness of a virtual environment by looking solely at the test subject and the way they mentally perceived the environment. The results from such methods should thus be solely based on the interpretation of the test subject and not the observer or system designer. The following subsections give a brief overview of the subjective presence measurement techniques that we have used as well as their application.

2.3.1.1 Quantitative presence measurement techniques

Quantitative data is particularly useful for statistical inference. Such data can tell with a high degree of confidence whether or not a virtual environment or any other application has been successful in what it is testing for.

Post-test Rating Scales / Questionnaires

Questionnaires provide a test subject with a set of questions that allow him/her to evaluate a virtual world based on their own interpretations. The most popular questionnaires are based on Likert scales and examples of such questionnaires are readily available. One of the more successful of these is the SUS questionnaire by Slater *et al.*[2]

Questionnaires provide concrete qualitative data representing a subjects feeling of presence in a virtual environment. However, the results of each questionnaire are still biased towards a particular individual's personality, therefore affecting each subject's view of presence.

2.3.1.2 Qualitative Methods

Quantitative data can be useful for certain applications, however, is often necessary to obtain qualitative data to get a more personal perspective on a system. For this reason, techniques such as focus groups and self reports can be instrumental in perfecting a system design.

3.2.1 Ethnographic Methods

Ethnographic studies are conducted by simply observing the subject while he/she is performing some task in a virtual environment. One key factor to such a study is that the subject not be bothered during the observation, or better still, that the subject not even know that he/she is being watched.

2.3.2 Objective Presence Measurement Techniques

Objective presence measurement techniques aim to view presence in exactly the opposite light to the way that subjective techniques do. These techniques do not rely on an individual's testimony of the success or failure of the virtual environment but rather looks at concrete evidence that might suggest whether an individual feels present or not in such an environment. The Collins English dictionary describes the word objective in exactly this way as follows:

> *undistorted by emotion or personal bias*
>
> *of or relating to actual and external phenomena as opposed to thoughts, feelings etc.*

The subsections to follow present some techniques that can be used to objectively measure a subjects feeling of presence in a virtual environment.

2.3.2.1 Qualitative Presence Measurement Techniques

Social Responses

Social behaviour such as smiling, laughing, gestures, body and head movements, eye contact, vocal cues etc. can be used to evaluate a subjects feeling of presence in a virtual environment. Slater [9] exploited these techniques by developing a virtual public speaking environment in which subjects were instructed to talk to a virtual audience for five minutes. The attitude of the audience varied from interested to bored to rude and showed results relating to self-rating fall as the audience's attitude changed for the worse. His experiment proved that all the above social acts have an effect on a test subjects level of presence.

Postural Responses

IJsselstein *et al.* [1] discuss how certain changes in posture can be an indication that a subject feels present in an environment. The argument is that if the virtual world feels real enough, the subject can experience vection. Such postural responses show that these subjects feel a certain degree of presence in such environments.

2.4 EXPERIMENTATION METHODS & PERFORMANCE MEASUREMENT

A major component to the effective testing of the optical tracking system is performance measurement. Arthur *et al* [10] describes a method whereby reactions times are recorded as part of the application in order to determine the speed at which a test subject can perform a certain task under varying conditions. These times can be used to draw valuable conclusions about the system and the ways in which it can be improved.

Variable Accuracy Performance:

Howlett *et al.* [11] mention an experimental evaluation test whereby experiments are conducted at different accuracies and test subjects are compelled to choose the level of accuracy that they deem to be adequate. This test could be modified slightly to fit the needs of the optical tracking system where the accuracy is determined by the level of sensitivity of the system. In other words, less sensitivity will result in a noticeable lag of the environment and too much sensitivity will result in an exaggerated change to the environment.

<MR> where M and R are four bits each.

M is the four bits which contain information about movement and R contain information about rotation.

Then M = <UDLR> and R = <UDLR> where for both cases U,D,L and R are one bit pertaining to Up, Down, Left and Right respectively.

If such a bit equals one, then it is assumed that motion of either a movement or a rotation is performed by the user.

Walkthrough Technique:

The walkthrough technique requires a test subject to walk through a virtual environment by using a human-computer interaction interface. This technique enables a test subject to experience the presence facilitating effects such as 3D sound or stereo vision while providing the experimenter with valuable data. This data can the be used to prove or disprove the initial hypotheses.

3. METHOD

3.1 Marker Correspondence

Matching up corresponding markers from each camera is achieved using knowledge of the marker pattern and simplifying assumptions about the user's pose based on the fact that the system is to be implemented in a Fish Tank VR environment. Consequently, a user cannot turn his/her head away from the screen without losing sight of it.

It is assumed that the user does not rotate more than 90 degrees from centre about any axis, so that he/she is always facing the screen and is not upside-down. These assumptions do not place unreasonable restrictions on users' movements and greatly assisted in implementing a simple prototype quickly.

These assumptions allow the system to identify the top, bottom left and bottom right markers uniquely by firstly registering the marker found with the greatest y-value as the top marker. The assumptions then dictate that of the remaining markers found, the one with the greatest x-value corresponds to the right one.

3.2 Motion Tracking

Motion tracking was designed and implemented to make use of the informational data from the previous time step as well as the current time step. The difference in this data would then correspond to the movement performed by the user.

3.2.1 Using the stereo camera setup of the VRRoom

Also, because of the physical constraints of the VRRoom and the field of view of the cameras, it was decided to implement motion tracking as follows. There exists a range known as the dead zone, this range constitutes no movement. When the user is in this range no movement takes place and if he was moving in the environment all movement stops. If the user then leaves this range, the movement being sent to update the environment corresponds to the point of exit of the range. If the user walks forward to exit the range, the environment is indefinitely update with the walk forward command until the user return to inside the dead zone.

3.2.2 Using the single camera for the ARToolKit Desktop solution

Again a dead zone was designated, however for this solution movement inside the dead zone does results in updates being sent to the environment. For this solution turning left and right as well as looking up and down were also implemented. For looking up and down no dead zone was designated, the reason for this is that in the environment a person can also only look up or down by a 90 degree angle max. The turning left and turning right features made use of the dead zone, in that when the total angle being turned away from the initial is greater than a specified threshold, indefinite turning commands would be sent to the environment. Indefinite motion only stopped once the user enters the dead zone again. Motion inside the dead zone for the ARToolKit implementation yielded corresponding motion updates for the environment.

3.2.3 Communicating the motion data to virtual environments

Communication between the motion tracking module and the virtual environmentswas implemented as a one way windows named pipe. This connection was responsible for sending the correct motion data from the motion tracking facility to the user environments. A C++ STL feature known as bitset was used as the carrier of the data. A bitset was chosen for efficiency as well as being able to send multiple types of movement in a single command. The encoded format of the bitset is explained by the following:

3.3 Experimental Design

An optical tracking system is intended for use in a variety of applications. These applications differ from virtual environment to virtual environment, so we have had to cater for all such variable conditions. Various techniques can be used to test both virtual and non-virtual environments. We feel, however, that since a tracking system can be used for so many different purposes, it is necessary to combine virtual environment testing techniques as well as non-virtual testing techniques. To this end we have decided to use three main experimental methods. These methods are listed below and a short overview of these techniques will then be given.

- Walkthrough (obstacle avoidance) / Navigation technique (Virtual)
- Modified Sensitivity Test / Accuracy Experiment (Non-virtual)

3.3.1 Walkthrough Technique

The first technique we have chosen to use is the walkthrough technique, discussed in Section 2.4. This technique enables us to determine whether or not the optical tracking system is intuitive or not to use. In this experiment, the test subject is instructed to walk through an environment and pick up objects along the way. These objects not only give the test subject a task to fulfill, but also serve as route markers that can guide the test subject on a pre-determined path. This path has been carefully chosen to test the optical tracking system's ability to handle clear straight runs as well as tight corner navigation. Figure 6 was taken from the start of the actual environment where the pickup can clearly be seen.

Figure 2 : Screenshot of pickup visible at the start of the walkthrough task

In our particular implementation of this technique, the time that it took a test subject to complete a pickup was recorded. In other words, times were taken between the starting position and the first object that required picking up, thus recording the time taken to pick up the first object. The timer was then reset and the time taken to pick up the second object from the position of the first object was recorded, and so on.

This test was performed twice as part of our experimental process. The rationale behind this was to account for the effects of learning in order to determine whether or not the system becomes easier to use over time. Although the time between the first and second walkthrough was relatively short, it was still possible to pick up on these effects. Another important aspect of the walkthrough was that test subjects experienced full 3D sound as well as stereo vision while completing the task. The use of the tracking system as well as these effects aimed to provide an experience that lead to some degree of presence.

4.3.2 Sensitivity Test / Modified Sensitivity Test

The second technique which we decided upon is the modified sensitivity test. This test was also discussed Section 2.4 and was used by Howlett *et al* [11] to conduct their experiments.

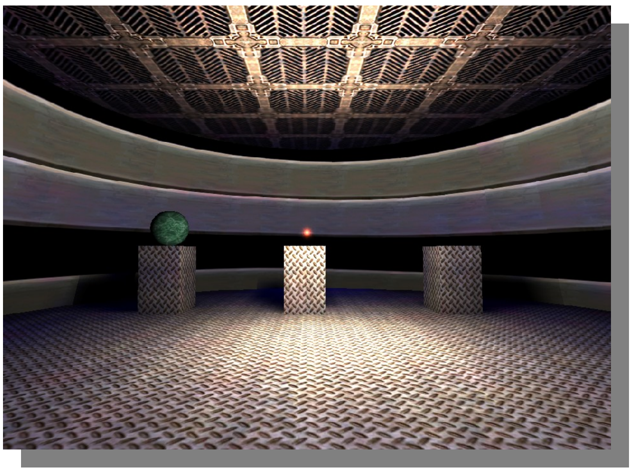

Figure 3 : Screenshot of the sensitivity task where green pickup can be seen

The modified sensitivity test is an extension to the simple sensitivity test that provides a mechanism by which to determine the speed at which the system can be used. This experiment required of the test subjects to select a number of objects from the scene. However, unlike the simple sensitivity test, this test involved a timing mechanism, thereby allowing a user to select an object only in the allocated amount of time. This time decreased over the course of the experiment, and inevitably lead to the objects remaining visible for only a few seconds. To make this task even more challenging, the size of the objects decreased over time thereby making the selection area smaller and smaller. In order to complete this test, test subjects had to select 25 objects, twice, thereby requiring them to select 50 objects in total. This gave a test subject the opportunity to try to select an object that was missed out on the previous round. Test subjects are, however, not conscious of the fact that this test gets repeated and are under the impression that only a single test is being performed. From this test we showed that the system can operate in a highly paced environment, as well as be accurate enough to complete the task.

Due to practical and implementation issues, the full optical tracking system was not implemented. The experimental design discussed above is only concerned with the experiments conducted using the keyboard and mouse.

4. RESULTS

4.1 Latency

The latency of the system is simply the time it takes for a user's motion to be updated on the display. Of course, this cannot be measured directly, so each time the camera frames were captured, a timestamp was added to the data. When the Sensor module had finished all its processing, a second timestamp was measured and the difference between these timestamps was taken as the latency. The timestamps were measured using the Windows high resolution timer functions, QueryPerformanceCounter() and QueryPerformanceFrequency(). This timer counts the number of clock ticks since the processor started. The frequency of the timer on the system used was 50000Hz which provides better than 1 millisecond accuracy. Since the order of magnitude of the latency is greater than 1 millisecond, this allows for accurate timing across a multi-process application.

Testing the system on a 3.4Ghz processor with 1Gb RAM, the average latency was found to be about 10 milliseconds. The time taken to process a frame once it had been received by the Sensor module was found to be about 6 milliseconds.

These results seem to be very good. Unfortunately, this seemingly low latency could not be reconciled with true latency. The true latency was often easily noticeable, which means that it must have been around 100 milliseconds. No data point ever exceeded 13 milliseconds. So, it is most probable that, while the DirectShow drivers can deliver images with such a high frequency, the images aren't being updated with that same frequency. Since the cameras are supposed to run at a maximum of 30 Hz, the true latency must be at least 33 milliseconds excluding the time taken to process the images. Thus, a lower bound for the true latency is 40 milliseconds, which includes the 6 milliseconds required to process the images.

For more accurate timing of the true latency of the system, one would need a method of determining when an image has been updated

4.2 Depth accuracy of the ARToolKit

To measure the accuracy of the depth report by the system, the marker pattern was kept at a constant distance of 300mm and 600mm away from the camera. Table 2, below, shows the measured information for depth at these distances, over a period of 20 time steps:

Time step	Actual Dist. (mm)	Measured Depth (mm)	Actual Dist. (mm)	Measured Depth (mm)
1	300	222.412	600	435.212
2	300	222.932	600	434.291
3	300	224.391	600	433.809
4	300	225.131	600	433.278
5	300	222.612	600	433.229
6	300	220.962	600	433.015
7	300	219.149	600	433.399
8	300	218.26	600	433.836
9	300	218.797	600	434.791
10	300	218.39	600	435.027
11	300	216.052	600	434.948
12	300	215.856	600	434.831
13	300	215.821	600	435.295
14	300	216.63	600	435.387
15	300	216.771	600	435.930
16	300	217.48	600	435.675
17	300	217.802	600	436.108
18	300	218.926	600	435.183
19	300	219.673	600	435.771
20	300	219.462	600	436.508
AVG	**300**	**219.375**	**600**	**434.776**

The information captured in Table 2, above, shows that there exists a jitter in the depth information. This jitter can be explained by attributing it to approximation errors introduced during the calculation of the scale factor in template matching. Another factor which could lead to jitter is the refresh rate of the camera as well as not having fully calibrated the camera before capturing the depth information.

From the results it can be seen that the measured depth distances differs from the actual values of the distance. However the ratio of the actual : measured distance remains relatively constant during the transition from the first position of 300mm away from the camera to 600mm away from the camera. This ratio being obtained as, 1:1.37 with an uncertainty of ±0.02 in the average depth distance.

The following factors could be attributed to this discrepancy between measured and actual depth distance.

- Improper calibration of the camera
- The scale factor of the marker used

For the first of these, if the camera is not calibrated correctly it might pick the marker up at a different scale then is actually used. If the scale of the marker is different to that of the template marker, the template matching code would then have to take into account an extra scaling factor. This extra scaling factor is however not taken into account in the ARToolKit, which would then lead to the discrepancy between the measured and the actual distances. The ratio of this discrepancy would then be related to the ratio of the scale of the template to the scale of the marker being used.

4.3 Experimental Results

In order to determine whether our optical tracking system is able to work as well as conventional user-computer interaction methodologies, we conducted user testing by using the keyboard and mouse. The data obtained from these experiments is used as a benchmark with which to compare the data obtained from the optical tracking system experiments. It is important to note that test subjects have been classified into groups depending on their level of experience. These groups are as follows:

- Non-gaming (NG)
  Rarely to never play games
- Non-First Person Shooter (NFPS)
  Does play games but not FPS games
- First Person Shooter (FPS) Plays primarily FPS games

4.3.1 Walkthrough Results:

Speed:

This section will provide a short summary of the results obtained from the user experiments conducted using the keyboard and mouse as well as provide a discussion on these results.

The results from the walkthrough experiment showed a direct correlation between experience and the time taken to complete the task. Those test subjects that indicated that they played first person shooter games had times substantially less than their non-gaming counterparts. These times ranged from 55.50 seconds for avid game players to 365.50 seconds for the non-game players. The poorer times indicate a distinct lack of keyboard and mouse coordination skills which was also evident from the ethnographic observations made. Test subjects that performed poorly found it difficult to navigate through the environment and occasionally got disorientated when looking straight down at the floor or straight up at a ceiling. Some test subjects had to be assisted on occasion as they could not orientate themselves again after losing control of the keyboard and mouse.

As discussed in Section 3.3.1, the walkthrough test was completed twice to take learning effects into account. The results of the second test showed an average decrease in the time taken to complete the task. This decrease could be attributed to the test subjects remembering the path to the end or simply due to a higher confidence in the use of the keyboard and mouse. Test subjects familiar with first person shooting games showed a much lower average decrease in time from the first attempt to the second, indicating that they are familiar with the keyboard and mouse interface. This shows that the learning of the path to the end of the task played only a minor role in the decrease in times of test subjects unfamiliar with this interface. This was particularly evident in the case of one of our test subjects who completed the second task quicker by two and a half minutes. This test subject had no coordination skills during the first attempt and slowly learned how to use the first person navigation interface. Although her time was substantially higher than the gaming enthusiasts, this large time decrease shows how quickly the keyboard and mouse interface can be learned in order to complete a task.

As predicted, the time taken to complete the walkthrough task was approximately 97.46 seconds on the second attempt, whereas the average time on the first attempt was 146.10 seconds.

4.3.2 Modified Sensitivity Test

Speed:

Since the modified sensitivity experiment was conducted twice, although, as mentioned in Section 4.3.2, test subjects were unaware of this fact, we were able to find the average times taken to select all objects on each of the two rounds. We found that the average time to select the second group of 25 objects was on average 10.18% quicker than that of the first round, with an average selection speed of 1.251 seconds per object, see **Error! Reference source not found.**. This clearly shows that test subjects got familiar enough with the mouse interface and scroll speed to better their time within a couple of minutes of starting the task. Once again, this proves that the mouse interface does not take much practice to learn how to use.

In terms of experience levels, test subjects that indicated that they were avid gamers performed much better than their counterparts in terms of selecting the objects quicker. Such test subjects registered an average selection time of 818.37ms compared to 1269.24ms for their non-gaming counterparts.

Accuracy:

The results obtained from our experiments showed that over the course of the two rounds of the modified sensitivity test, test subjects were able to pick up all the objects in the test. In other words, if in round one, a test subject missed object number four, they were able to pick up this object in round two. Since no movement was required for this test, only the mouse had to be used for object selection. Non-gaming test subjects, even though their selection times were slower, could manage to select all the objects from the environment. This could be attributed to the fact that no keyboard and mouse coordination was required, and that all test subjects were familiar with the use of a mouse by itself. This test showed that the mouse interface was suitable for use in a highly paced environment that required a high degree of accuracy.

Usability:

In order to determine whether the keyboard and mouse interface has a suitable degree of usability, three questions were added to the end of the questionnaire that was completed by all test subjects. These questions asked of test subjects to state what they liked about the system, what they disliked about the system and asked for any improvements that they through could enhance the experience. In terms of usability, the only problem that could be found with the keyboard and mouse interface was that the use of wireless input devices, in our case Bluetooth, was too slow and that wired optical mice could provide better performance. These aspects are, however, practical issues that do not detract from the use of a keyboard and mouse as a human-computer interaction solution. In other words, by providing users with wired mice, the use of the keyboard and mouse would be an attractive interface to complete tasks such as those presented during the experiment.

4.3.3 Presence Issues:

Since the measurement of presence is a subjective measure, test subjects were asked to rate the environments by completing the Temple presence questionnaire.

A lot of attention was paid to creating a realistic looking 3D environment, particularly for the walkthrough task. As already discussed, stereoscopic vision as well as 3D sound was incorporated into this task to fulfill these requirements. The results obtained from the questionnaires show that the test subjects indeed felt a sense of depth, as an average score of 71.43% was indicated for objects appearing to be real. In a broader sense, the results show that test subjects felt spatially present 67.86% of the time. This result is lower than anticipated perhaps due to the fact that the environment looked more like a cartoon than a real place. However, the results show that the stereoscopic vision did its part to enhancing presence as anticipated.

Test subjects felt mentally immersed on average 66.37% of the time. This value is as expected since the environments seemed reasonably realistic to test subjects and made them feel as though they were in the environment. This result has a direct correlation to the realism factor of 71.43% reported in the previous paragraph. In terms of perceptual realism, test subjects indicated a 42.14% realism scale. Since the environments contained few objects and no virtual people, we were not surprised by these poor results.

The final set of questions in the questionnaire are taken from the NASA Task Load Index as described in Section 4.8.2. According to the results obtained, the tasks set only required a 32.65% concentration level, differing substantially from the 66.37% recorded above. However, the latter has to do with mental immersion whereas the former has to do with mentally challenging tasks. In other words, test subjects did not find the tasks challenging but did feel mentally immersed in the environment. A low score was anticipated because of the familiarity of all test subjects with the keyboard and mouse. An interesting observation was made once again with respect to experience levels. Test subjects that indicated that they were not avid gamers and took a longer period of time to complete the walkthrough task, rated the metal demand aspect higher than those test subjects who play first person shooter games regularly. These results show that it is imperative that an input device be easy to use in order for users to focus on the task at hand and not the interface.

Since the tasks resemble game-like environments, many test subjects could complete the tasks without much concentration. For the optical tracking system, however, we expect that tests subjects will have to concentrate more as the interface to the environment is unfamiliar. Similarly, test subjects found that the tasks were not physically challenging, giving this aspect a rating of 26.53%. Once again, these results are as expected as the keyboard and mouse do not require much physical ability to operate. We do expect, however, that this rating will increase during the optical tracking experiments as they will require a more physical approach to operate.

5. CONCLUSION

The final optical tracking system is far from finished. Various complications in terms of technology and time have arisen during the course of this project. The failure to produce a working system can be attributed partly to the following:

- Cameras were obtained too late into the project due to stock shortages
- Lack of communication between group members responsible for the development of layers one and two, see Figure 1, lead to double work being done
- Integration of modules occurred too late leading to unforeseen problems
- Mounting of cameras too late lead to a number of unforeseen problems
- Unclear understanding as to the capabilities and shortcomings of the toolkits used

To date, our group has managed to implement a system that works in low-light conditions and has a one degree of freedom movement capability. Although this system does not meet the six degree of freedom requirement, it has, however, shown that it is possible in practice to use low-cost cameras in order to build an optical tracking system. The LED head mounted system that was developed has been tested and works well for the application which we have implemented, although there is room for improvement. Although a separate six degree of freedom system was partially developed and could be used for the walkthrough experiment, discussed in Section 4.3.1, this system did not work in low-light conditions, was exceptionally difficult to control and did not use a head mounted LED detection system. User testing was therefore not conducted due to these severe shortcomings.

In terms of our initial hypotheses listed in Section 2.2, hypotheses numbered one to four were the responsibility of the other group members' sections. Although the system was not completed, it was still possible to prove all four of these hypotheses true. Unfortunately, due to the prototypical state of the system and its lack of functionality, it was not possible to conduct user testing on the optical tracking system. It is therefore unfortunate that hypotheses numbered five through seven could not be proven.

In terms of requirements, the table below summarises the progress made to date:

1	The system should be low-cost	✓
2	The system should work in low-light conditions	✓
3	The system should provide movement capabilities for six degrees of freedom	×
4	The system should be easy to use	×
5	The system should be efficient enough to minimize lag and delay	×
6	The system should have a moderate to high degree of accuracy	×

Table 1 : Requirements met at the end of the project

The crosses indicate that the requirements could not be met with the current implementation or that testing could not be carried out to verify these requirements. We hope to perform these tests in the near future once we have the full system implemented.

6. FUTURE WORK

Extending the marker matching algorithm to detect more than 3 markers and to match up markers at arbitrary orientations is of primary importance. This could be achieved using epipolar geometry or template matching. This extension would allow for full 6DOF tracking. Of course, different hardware would need to be used to allow the markers to remain visible at all times.

For the user and system testing component of the project, much future work still needs to be done as user testing using the optical tracking system could not be performed. This work would consist of running the experiments that have been designed exclusively for the optical tracking system as well as analysing the results in order to determine whether the hypotheses set out in Section 2.2 have been proven or not. Since our tests have been conducted using only eight test subjects, future work should aim to increase the amount of test subjects in order to obtain better results.

One of the main sections of the optical tracking software that has not yet been implemented is motion prediction using Kalman filtering. Possible future work would include the creation of an experiment to test the effects of navigation and task completion in virtual environments with and without Kalman filtering. Results from such experiments should prove that an optical tracking system can be made more usable and lead to decreased task completion times if Kalman filtering is applied.

7. REFERENCES

[1] Wijnand A. IJsselsteijn, Huib de Ridder, Jonathan Freeman, and S.E. Avons. *Presence: Concept, determinants and measurement.*

[2] Slater, M., Usoh, M., & Steed, A. (1994). *Depth of presence in virtual environments.* Presence: Teleoperators and Virtual Environments, 3, 130-144.

[3] Bob G. Witmer and Michael J. Singer. *Measuring Presence in virtual Environments: A Presence Questionnaire.* Presence, Vol. 7, No. 3, June 1998, 225–240

[4] Retaux, X. (2003). *Presence in the environment: theories, methodologies and applications to video games.* Psychology Journal, 1(3), 284 - 310.

[5] McGreevy, M. W. (1992). *The presence of field geologists in Mars-like terrain.* Presence: Teleoperators and Virtual Environments, 1, 375 - 403.

[6] Joy van Baren and Wijnand IJsselsteijn. www.presence-research.org/Qualitative.html#Thinking

[7] Freeman, J. & Avons, S.E. (2000). *Focus group exploration of presence through advanced broadcast services.* Proceedings of the SPIE, 3959, 530-164.

[8] Turner, S., Turner, P., Caroll, F., O'Neill, S., Benyon, D., McCall, R., et al. (2003). *Re-creating the Botanics: Towards a sense of place in virtual environments.* Paper presented at the 3rd UK Environmental Psychology Conference, Aberdeen, 23-25 June 2003.

[9] Slater, M. (1999b). *Co-presence as an amplifier of emotion.* Invited talk at the 2nd International Workshop on Presence, University of Essex, Colchester, 6-7 April 1999.

[10] Arthur, K.W, Booth, K. S *et al*. *Evaluating 3D Task Performance for Fish Tank Virtual Worlds.* ACM Transactions on Information Systems. Vol. 11, No. 3, July 1993. Pages 239-265

[11] Howlett, S *et al*. *A framework for comparing task performance in real and virtual scenes.* Trinity College.

[12] J.D. Mulder and R. van Liere. *Enhancing fish tank VR*. In S. Feiner and D. Thaimann, editors, Proceedings of the IEEE Virtual Reality 2000.

[13] R. van Liere and A. van Rhijn. *Search Space Reduction in Optical Tracking*. In Proceedings of the workshop on Virtual environments 2003 EGVE '03.

[14] M.R. Shortis, T.A. Clarke and T. Short. *A comparison of some techniques for the subpixel location of discrete target images*. In SPIE Videometrics III. 1994.

[15] J.D. Mulder, J. Jansen and A. van Rhijn. *An Affordable Optical Head Tracking System for Desktop VR/AR Systems.* Proceedings of the workshop on Virtual environments 2003.