PUBLICATION P5

Ilmonen, Tommi and Reunanen, Markku Virtual Pockets in Virtual Reality. Virtual Environments 2005 Eurographics/ACM SIGGRAPH Symposium Proceedings, pages 163–170, 2005.

IPT & EGVE Workshop (2005)
R. Blach, E. Kjems (Editors)

Virtual Pockets in Virtual Reality

Tommi Ilmonen & Markku Reunanen
Helsinki University of Technology
Telecommunications Software and Multimedia Laboratory
tilmonen,marq@tml.hut.fi

Abstract

This paper presents the use of the virtual pocket user interface metaphor in virtual environments (VEs). Virtual pockets act like real-world pockets — the user can put tools into pockets and select them with simple gestures. The pockets are body-centric and they can be used while the user is focusing on some object in the VE. We present the behavior of our virtual pocket implementation and compare its performance with different variations. The user tests show that virtual pockets are a useful metaphor and that selection of the correct feedback method has a strong impact on how well the pockets work.

Categories and Subject Descriptors (according to ACM CCS):
H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities
H.5.2 [User Interfaces]: Interaction styles

1. Introduction

Tool selection and configuration are common tasks in most computer applications. In desktop systems with graphical user interfaces menus and buttons are effectively used for this task, while in virtual/augmented reality systems few standards exist. The virtual pocket metaphor is one way to handle the tool selection needs.

Typical VE approach to tool selection has been to include 3D signs/widgets into the world. These are problematic since they can be occluded by other geometry, they may occlude other geometry, they may be in a difficult position and they may disturb by their very presence. During the past years alternative solutions to this problem have been proposed.

The novelty of our work is in introducing the virtual pocket metaphor to virtual environments and the results from the usability tests that we have conducted.

Over the years a number of techniques have been used for tool selection in VR. A comprehensive list of different strategies can be found in Weshce's work [Wes03]. Wesche also created a new tool selection metaphor — the ToolFinger. Mine has presented physical mnemonics that are similar to virtual pockets [MFPBS97]. In his work widgets were located either to the user's hands or out of sight (relative to head or body). Research on VR interaction has been carried out by e.g. Conner [CSH*92].

Virtual pockets were first introduced by Lehikoinen [Leh01], who used them in conjunction with wearable computers. Compared to Lehikoinen our implementation uses slightly different logic and we use the approach in VR. We have developed and used the virtual pocket system in a 3D drawing application. This application is similar to previous work by Keefe [KFM*01] and Mäkelä [MRTI04].

2. The Application

The virtual pockets are used as part of a 3D drawing software. It is an in-house application that enables artists to draw 3D shapes in a CAVE-like environment (see Figure 1). The user has a selection of drawing primitives (polygon tube, line, particle clouds) and tools to modify the graphics (magnet, eraser, rotation etc.). The tools are used in real-time with hand gestures.

Each tool has parameters that can be adjusted for the situation. In the past all the tool selection and configuration was done via "kiosk" — a large collection of 3D widgets (see Figure 2) — that was used for all application control. The

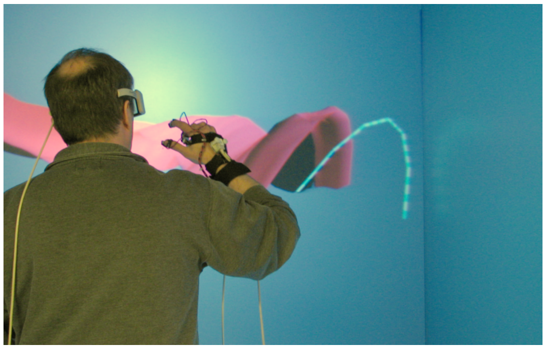


Figure 1: *An artist is drawing with a tube. The drawing cursor is presented as an arc.*

kiosk is a powerful control center but also a source of many problems: The drawing act is interrupted whenever the artist needs to change or adjust a tool. Also the drawing is either replaced or hidden by the large kiosk. This makes tool configuration difficult, for example when trying to match the color of a new drawing tool to existing graphics.

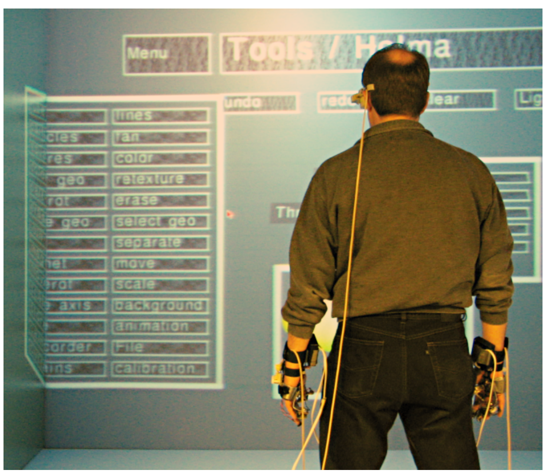


Figure 2: *Tool selection with the kiosk.*

These concerns led us to develop a quick edit mode (or a sub-kiosk). In this mode only the most important parameters of the active tool were brought to the virtual environment. The tools could be placed with the drawing hand. This improved the tool configuration since the tool could be adjusted without interrupting the work-flow completely.

Even so the users desired methods to change the tool without changing posture or breaking eye contact with the object that was manipulated. Virtual pockets were seen as a method to achieve this goal. We also integrated some features of the earlier quick edit mode into the virtual pockets.

2.1. Environment and Devices

We have a normal Cave-like virtual reality installation. This system has four back-projected walls with active stereo system, 14 loudspeakers for surround sound and a magnetic motion tracking system. The setup can be seen in Figures 4 and 1. The users' fingers are are tracked with a Näprä (see Figure 3) [RPIM05]. This is a custom device that can track the location and contact of the first three fingers. It also has a two-way button for the ring finger and a one-way button for the little finger.

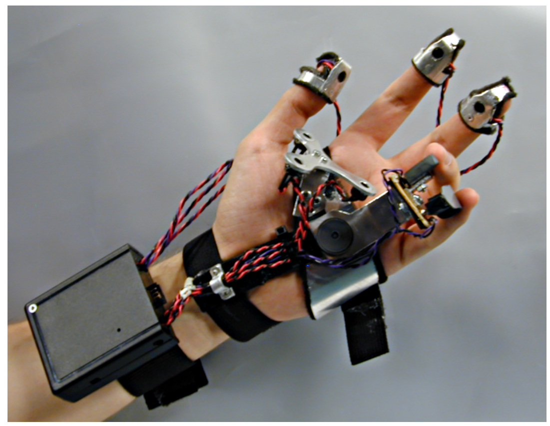


Figure 3: *The Näprä finger-tip tracking device.*

The graphics are generated with an SGI Onyx2 system. In these tests the frame rate was typically 20-30 Hz.

3. The Virtual Pocket Metaphor

The virtual pocket metaphor mimics real-world pockets. A user can put tools into pockets, take them into use from the pocket. The system relies on human spatial memory — for example an electrician remembers the pocket where he keeps a particular screwdriver. Unlike real-world pockets a virtual pocket can contain only a single tool at a time. The user can perform three tasks with a pocket:

1. By placing his/her hand inside the pocket and touching thumb with either index- or middle finger the user can activate the tool that is located in that pocket. This enables the user to change tool quickly, while maintaining focus in the work.
2. If the user maintains contact between fingers and pulls his/her hand out of the pocket then widgets that are related to the tool become visible. The widgets move with the hand until the user releases contact between the fingers. This way the user can bring the tool's widgets into the work area and tune the tool. The widgets can be placed so that they are not occluded by other geometry.

3. Besides adjusting the tool parameters the user can also change the pocket's tool with the widgets. Any change in tool or its parameters is automatically stored into the relevant virtual pocket. This way the users can configure the tool-set to suit the needs of the current task.

These features together enable the following work-flow: Pocket activation (preview) → tool selection → tool adjustment → tool use. The adjustment phase can be skipped if it is not necessary.

Feedback affects how easy it is to use the pockets. We have tested the following feedback methods:

1. No feedback. (N)
2. Visual feedback in the form of an area display. This method displays the pocket areas and their activation when the user's hand is inside any of the pockets. The display is shown relative to user's head — it is always in the middle of the field of vision. The visual feedback is displayed in figure 4. (V)
3. Audio feedback when a tool is activated. When this method is used the areas emit tool-specific sounds that indicate that user's hand is inside some pocket and as the tool is selected. We took the audio icons from the open-source KDE-application suite. (A)
4. Combination of the plain visual feedback and audio feedback. (AV)
5. Enhanced visual feedback with tool texts on top of the area display. This method adds texts on top of the area display. (VT)

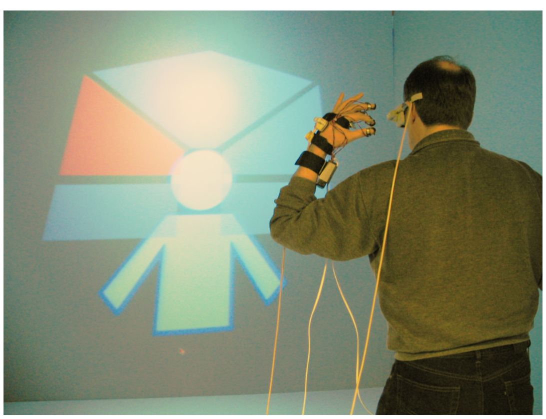


Figure 4: *Visual area display in an otherwise empty world. The first pocket is on top the left shoulder and fifth on top of the right shoulder (for right-handed users).*

Further variations are also possible, but we decided to limit our study to these choices. All of these methods have some inherent strengths and weaknesses that depend on the use context. The amount of other visual and audio activity can make some feedback method particularly good or inappropriate.

In addition to the explicit feedback the application provides implicit feedback. For example when a tool is changed the visual shape of the user's 3D cursor is altered to match the new tool. Often users can utilize this feedback to determine whether that they have changed the tool successfully. Since some tools share the same 3D cursor this information does not always offer sufficient feedback.

In principle the pockets could be located anywhere around user's body. There are some practical constraints to this freedom. 1) The user should not put his/her hands into the pockets accidentally. This implies that the pockets should not be in the hands' working area. 2) We'd like to use minimum number of motion tracking sensors due to user comfort and to minimize system complexity. Since we already had placed a sensor in the user's stereo glasses a natural solution was to place the pockets around the user's head. Since the application has complete knowledge about the head and hand positions we can reliably tell when the user's hand is in some pocket. Also the area around the head is seldom used when drawing with the system. The pocket locations where designed to work well for an average-length person. We assumed that the identical pocket locations could be used regardless of the physical size of the user. Figure 5 shows how the pockets were organized around the user's head in the test situation.

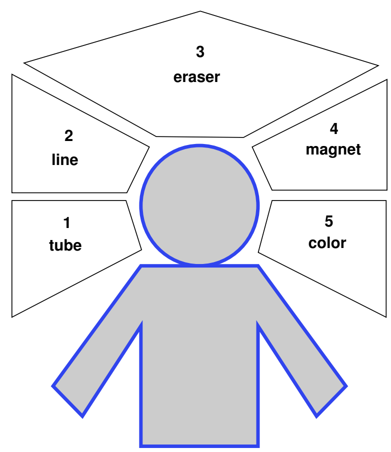


Figure 5: *The locations and contents of the pockets.*

4. User Tests

The main emphasis of our user tests was to determine the overall feasibility of the approach and how well the different feedback methods work. The drawing application was the context of these tests. In the tests we used five feedback setups. There were three people present: The user, an instructor who interacted with the user and a technical assistant who ran the software and made observations. Each test session was composed of four parts:

Figure 6: *The 3D sculptures drawn by one test participant, shown in the drawing order.*

In the beginning the user was told what the test was about and they could also test the devices and application shortly. In this phase we used plain visual feedback for all users.

In the second phase the users were given simple performance tests where they only had to click the specific pockets. The instructor spoke aloud the number of the pocket to be clicked and the user reacted as quickly as possible. The assistant pressed mouse buttons as the instructor finished the utterance. The application stored both the assistant's button presses and the users' behavior into a file for later analysis. From this data we could calculate the time it took to select and configure a tool, errors in the tool selection and overall user performance. The order in which the feedback methods were used and the pockets selected was randomized to prevent learning effects from disturbing the analysis. Each participant used all five feedback methods and performed ten pocket selections with each method, resulting in 50 selections per user. This synthetic performance test gives plenty of accurate data for analysis.

In the third phase the users were given step-by-step instructions to draw a simple 3D sculpture. Each user drew five sculptures: a stick-man, a seagull, a flower, a kite and a bicycle. The drawings of one participant are shown in Figure 6. Each sculpture required the use of all available tools. The sculptures were drawn in eight stages, with each stage requiring a tool change. With five tasks this makes 40 selections. As in the previous phase the user actions were stored for later analysis. Both the feedback methods and the sculptures were presented in a randomized order. This is a more realistic use case, that provides both quantitative and qualitative data.

The fourth part of the test was an interview where the instructor asked a series of questions from the user, for example asking them to rate the feedback methods from the best to the worst. Throughout the test the instructor and the assistant observed the user and wrote down interesting findings. The sessions lasted 55–70 minutes.

Before conducting the actual user tests we evaluated the test setup with a pilot test. The pilot test confirmed that the test setup was feasible and that it is possible to use five pockets. We were prepared to change the test setup after the first user test, but it was not necessary.

4.1. Measurement Data Management

When the instructor told the user to select a pocket or use some tool the user was to make one (correct) pocket selection. In practice this is was not the case. Users often made multiple selections before they were happy with the result. For example people could make an incorrect selection and then make the correct one. Sometimes people even selected the same pocket more than once because they were uncertain if the first selection was successful. Since we needed to condense these multiple selections into one time value per test point we selected only one of the values.

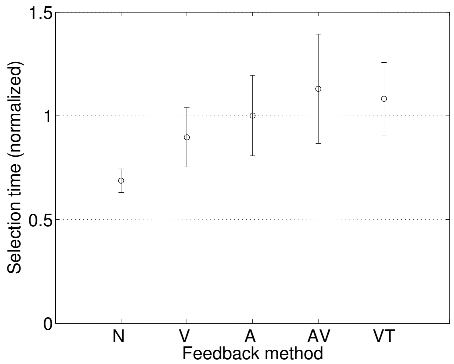
In the performance test we always used the last selection. This is the selection that the users were happy with. In the drawing test the situation was more complicated. Often a user could make the correct selection, then realize that they needed to change tool parameter (for example the drawing color) and finally re-select the tool and drag its parameters into the working area. In cases like this we would get a number of selection events before the user was happy with the result. Since we were testing the selection speed only we accepted the first correct selection. If there were multiple selections made quickly (within 0.5 second interval) we chose the last one since that is most probably the point when the user felt he/she had made the selection.

4.2. Test User Demographics

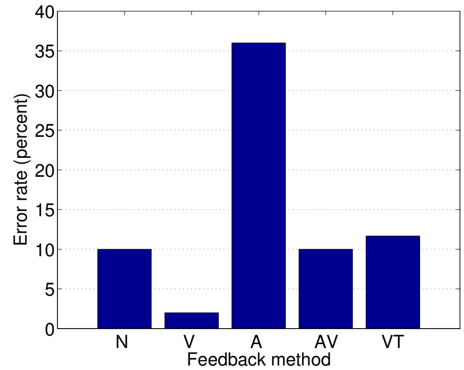
The test users were taken from our laboratory. Six men and one woman participated in the test. All had strong technical background. Some had visual arts as a hobby. Of the seven test users three had not used any kind of immersive VR system before. Compared to the average population this is a skewed one, but it most likely matches the typical user base of VR systems rather well.

5. Test Results

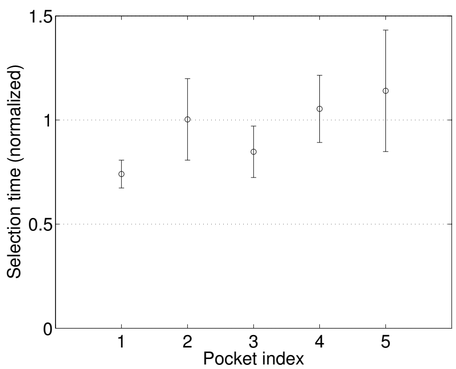
The test setup yielded both qualitative and quantitative data. The quantitative data is made up of the speed measurements



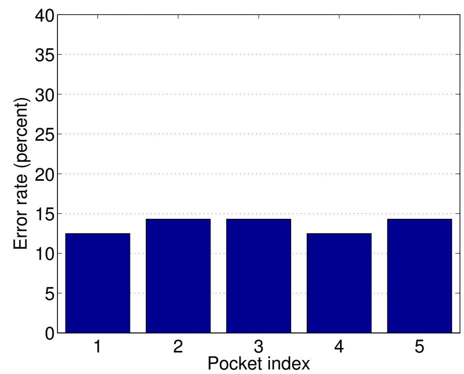
(a) Pocket access speeds with different feedback methods.



(b) Mean of error rates in the performance test with different feedback methods.



(c) The access speeds of different pockets.



(d) Mean of error rates in the performance test with different pockets.

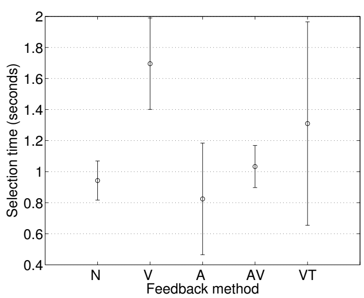
Figure 7: *Pocket selection speeds and error rates in the performance test.*

and error rate calculations. The qualitative results were obtained from user interviews and observations. We received a wealth of information about how the users liked the system and how they worked with the application. Keeping to the scope of this paper we only present findings that are directly related to the virtual pockets.

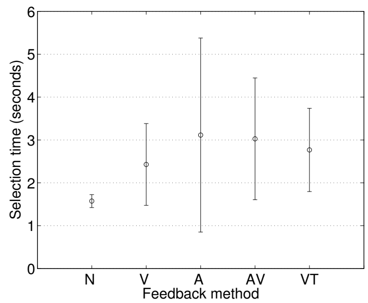
5.1. Quantitative Results

First we cover the simple performance tests. When analyzing the performance tests we dropped away the first test round from each user. This performance test data contains 40 data points per user, giving a total of 280 measurements. When combining the data from different users we normalized the operating speeds. This was done to counter the effect of having users with very different average speeds.

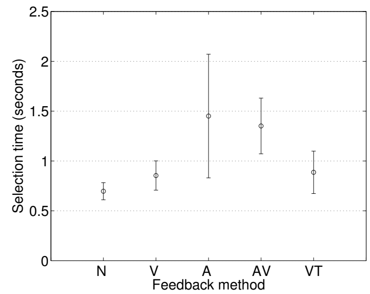
Operation speed is one of the most important aspects of the user interface. The selection times and their 95% confidence intervals have been collected into Figure 7(a). This graph reveals that the users performed fastest without any feedback. Adding more feedback made the selection slower, apparently because people would stop to evaluate the feedback before making the actual selection. As a statistical test method we used analysis of variance (ANOVA). Statistically one of methods differs from others, $F(4, 237)=4.2$, $p=0{,}0026$. According to a post-hoc test it is the no-feedback method differs from the others. The post-hoc test was Tukey's honestly significant difference criterion.



(a) The average selection speed and 95% confidence interval for Charlie. The tests were performed in the order: V–N–A–AV–VT.

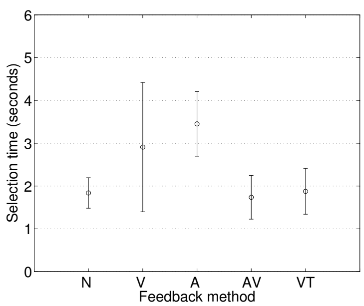


(b) The average selection speed and 95% confidence interval for John. The tests were performed in the order: VT–A–AV–V–N.

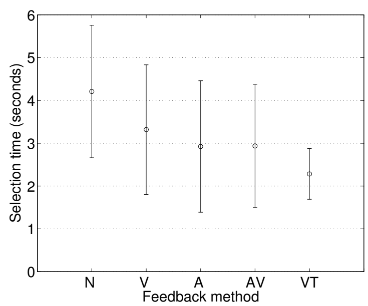


(c) The average selection speed and 95% confidence interval for Susan. The tests were performed in the order:V–VT–N–A–AV.

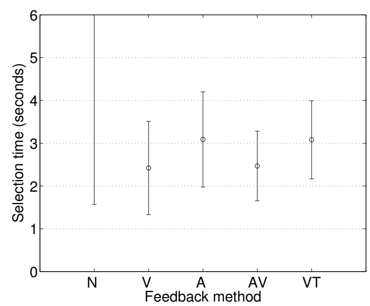
Figure 8: *The performance of the different feedback methods with different users in the performance test. Note the different Y-scales.*



(a) The average selection speed and 95% confidence interval for Charlie. The tests were performed in the order: A–V–VT–N–AV.



(b) The average selection speed and 95% confidence interval for John. The tests were performed in the order: V–N–A–AV–VT.



(c) The average selection speed and 95% confidence interval for Susan. The tests were performed in the order:A–N–AV–VT–V.

Figure 9: *The performance of the different feedback methods with different users when drawing.*

To highlight individual differences we have collected the performance timings of three very different users into Figure 8. Since learning takes place during the test these figures need to be read with caution.

The average error rates of the different feedback methods are collected into Figure 7(b). Surprisingly the plain visual feedback was clearly the most reliable, while combining it with other feedback made reliability worse. This implies that adding more feedback may confuse the user. The audio feedback was much worse than no feedback at all. It seems that audio feedbacks diverts the user's attention away from the actual selection task. It should also be noted that the users had used the plain visual feedback in the learning phase and were more familiar with it and the audio feedback was first introduced during performance tests.

We were also interested in knowing if some pockets would be more usable than others. Before running the analysis we mirrored the pocket indexing for the only left-handed user. The error rates in Figure 7(d) show that none of the pockets was particularly reliable or unreliable. The speed of operation was different however, with the first and third pocket clearly fastest and the fifth pocket slowest (see Figure 7(c)). In the first round most users had great trouble with the fifth pocket, but over time this effect diminishes. Statistically have different performance with F(4, 214)=4.97, p=0.0008.

A post-hoc test indicates that the fifth pocket differs from the first.

It should be noted that the average error rates do not tell the whole truth. In fact there were people who made hardly any mistakes with any feedback method while others made plenty with all the methods.

The drawing tests provided a different perspective to how the feedback methods compare. We have again removed the first drawing task from the analysis to minimize learning effects. Since each test consisted of eight steps we get 32 measurements per user. Few outliers were removed from the data set because either the assistant had made a mistake or we could not interpret the point where the user had actually made the selection.

In the drawing test users may make incorrect selections, but they must eventually make the correct selection to be able to go on with the test. In practice errors show up as a longer time to get to the correct tool. The relative performance of the feedback methods has been collected into Figure 10. The figure shows that for most people in this test visual feedback was necessary for efficient operation. Combining visual and other feedback seems to make little difference. The selection times are affected by the time it takes to make a selection and the time that is spent thinking about the upcoming task. Statistically one of methods differs from others with $F(4, 214)=4.97$, $p=0.0008$. A post-hoc test reveals that the N option differs from V, VT and AV,

An analysis of individual users' behavior reveals details that cannot be seen from the collective results. Three user's results have been collected into Figure 9.

Given the small number of test users we cannot say too much about how earlier VR experience affects the performance. In any case, both the fastest and slowest performers were people who had not used similar VR systems before.

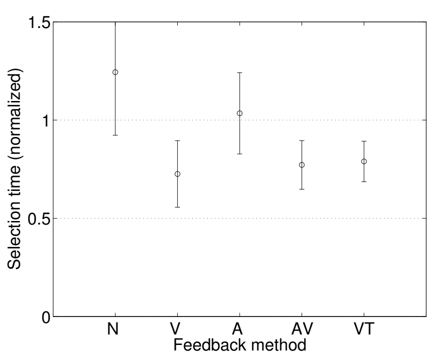


Figure 10: *The speed of the different feedback methods when drawing.*

Charlie	A	N	VT	AV	V
John	VT	A	AV	N	V
Wayne	AV	VT	V	A	N
Clint	VT	V	AV	A	N
Alfred	AV	VT	A	V	N
Susan	VT	AV	V	A	N
Luke	VT	AV	A	V	N

Table 1: *User preferences from the best to the worst. The names are not the real names of the users.*

5.2. Qualitative Results

Table 1 summarizes the user preferences obtained in these tests. We found out that the users had very different preferences for feedback. For example the audio and visual feedback methods were rated both high and low. This lack of concensus implies that the system feedback should be configurable to match different users' desires.

To give more depth to these figures we present some of the explanations that the users came up with when making the list of preferences. Charlie was annoyed by the graphical feedback. He experienced that it distracted him from the actual artistic work and came in the way. His hand motions frequently caused mis-triggerings — the system gave feedback even when he was not about to change any tool. He found the audio feedback pleasant since it used another modality (from graphics) to deliver information. He was the only one to rate the no feedback option into a favorable position.

Charlie and John were the only ones who rated the audio feedback to the first or second place. Both of them work with audio technology, implying that personal background may influence the preferences.

Clint had largely opposite views on the feedback methods. He found that the visual information was most reliable and the audio feedback was mostly annoying. Susan also rated the multi-feedback methods highest and she stated that "the more you have feedback the better it is". She was extremely accurate and made only 4 errors in the 50 test cases during the performance tests.

At any rate it can be seen that the multi-feedback methods (AV and VT) dominate user preferences while the no-feedback option is generally disliked. Plain visual feedback was rated lower than plain audio feedback, but when combined with text (VT) the visual feedback was superior to visual feedback with audio (AV). Almost all users had clear opinion of the best and the worst feedback method. The three in-between methods were not so easy to sort.

Learning is a factor that has a strong impact on the way the pockets are used. During the test each user became much more fluent in using the virtual pockets. With greater skill the role of the feedback diminishes. Many also commented that the preferred feedback method might change if they used the application for a longer time — the highly explicit visual feedback might become completely irrelevant after some time. The audio icons might also become annoying after prolonged use.

Fatigue was not a problem in these tests. All the users were explicitly offered pauses to avoid fatigue, but all refused the possibility. Only one of them found the experiments slightly tiring.

5.3. Discussion

All test users found the virtual pocket metaphor useful. The virtual pockets were easy to understand and they supported the work-flow when making the sculptures. We met most of our design goals: The pockets can be used without changing body posture and they do not interrupt the work. Going back to our design motives we can notice that the visual feedback conflicts with some of our original aims — the user cannot maintain steady eye contact with the art if the visual feedback suddenly appears in the middle of the field of vision. In these tests only Charlie was concerned about this topic.

If the user places his/her hands inside the pocket without meaning to change the tool the pockets should not be activated and feedback is not desired. There are situations when users do place their hands close to their head. The most common reason for this is when the users adjust the stereo glasses. The glasses are heavy and slide easily, needing frequent re-adjustment. Although this causes mis-activation it is not a serious problem since the user is anyhow occupied with adjusting the glasses and not about to draw anything. One user (Charlie) adopted a hand gesture that often caused mis-reaction.

We had assumed that our original pocket locations would fit all users. In practice we found out that some adjustment is necessary to optimize the system for different users. For example one very tall user would have liked larger pockets — for him the pockets were crammed just around the head. Users with short neck had trouble reaching the first and the fifth pocket — their shoulders took up pocket area and made pockets one and five difficult to reach.

Our tests measure how new users use the virtual pockets. While there seem to be large individual differences it is difficult to analyze them statistically due to the learning effects. There are also important variants that were not included in the tests. For example the users were not allowed to change the contents of the pockets. This is an important feature that probably affects the way feedback is utilized. Our tests do not address the issue of long-term usage. The methods might rate very differently after months or years of use. People might also develop more elaborate strategies for using the pockets.

Given the above results we will probably use visual and text feedback as the default, since it is most liked and has a low error-rate in real application usage. To meet different user needs we will also offer the possibility to configure the feedback.

6. Conclusions

We have shown that the virtual pocket metaphor works well in or application context in virtual reality and that the different feedback methods appeal to different users. There is no feedback method that would be objectively best. Both quantitative and qualitative results imply that the pocket system should be configurable to work well with different users. In future we will continue to work on making the implicit and explicit feedback work well together.

7. Acknowledgments

This work has been funded by Tekes, the National Technology Agency of Finland.

References

[CSH*92] CONNER B. D., SNIBBE S. S., HERNDON K. P., ROBBINS D. C., ZELEZNIK R. C., VAN DAM A.: Three-dimensional widgets. In *SI3D '92: Proceedings of the 1992 symposium on Interactive 3D graphics* (New York, NY, USA, 1992), ACM Press, pp. 183–188.

[KFM*01] KEEFE D., FELIZ D., MOSCOVICH T., LAIDLAW D., LAVIOLA J.: Cavepainting: a fully immersive 3d artistic medium and interactive experience. In *Proceedings on 2001 Symposium on Interactive 3D graphics* (2001), ACM Press New York, NY, USA, pp. 85–93.

[Leh01] LEHIKOINEN J.: Virtual pockets. In *Proceedings of the 5th World Multiconference on Systemics, Cybernetics and Informatics (SCI) 2001* (2001), vol. IV, pp. 479–484.

[MFPBS97] MINE M. R., FREDERICK P. BROOKS J., SEQUIN C. H.: Moving objects in space: exploiting proprioception in virtual-environment interaction. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1997), ACM Press/Addison-Wesley Publishing Co., pp. 19–26.

[MRTI04] MÄKELÄ W., REUNANEN M., TAKALA T., ILMONEN T.: Possibilities and limitations of immersive free-hand expression: a case study with professional artists. In *ACM Multimedia 2004 Conference* (October 2004), pp. 504–507.

[RPIM05] REUNANEN M., PALOVUORI K., ILMONEN T., MÄKELÄ W.: Näprä — affordable fingertip tracking with ultrasound. In *EGVE '05: Proceedings of the workshop on Virtual environments 2005* (2005).

[Wes03] WESCHE G.: The toolfinger: supporting complex direct manipulation in virtual environments. In *EGVE '03: Proceedings of the workshop on Virtual environments 2003* (New York, NY, USA, 2003), ACM Press, pp. 39–45.