# An Experimental Analysis of the Impact of Touch Screen Interaction Techniques for 3-D Positioning Tasks

Manuel Veit*  Antonio Capobianco†  Dominique Bechmann‡

University of Strasbourg – LSIIT (CNRS UMR 7005)

**ABSTRACT**

The use of Touch Screen Interaction (TSI) for 3-D interaction entails both the addition of new haptic cues and the separation of the manipulation of the Degrees of Freedom (DoF) of the task: a 3 DoF task must be transformed into a 2-D+1-D task to be completed using a touch screen. In this paper, we investigate the impact of these two factors in the context of a 3-D positioning task. Our goal is to identify their respective influence on subjective preferences and performance measurements. To that purpose, we conducted an experimental comparison of five positioning techniques, isolating the influence of each of these two factors. The results we obtained suggest that the addition of haptic cues does not influence the user precision. However, the decomposition of the task has a strong influence on accuracy. More precisely, separating the manipulation of the depth dimension leads to an increased precision while isolating other dimensions does not influence the results. To explain this result, we realised a behavioural analysis of the data. This study suggests that the differences in the performance may be linked to the perceptual structure of the techniques. A technique isolating the manipulation of the depth seems to have a more adapted perceptual structure than a technique separating the height, even if those two dimensions are equally involved in the realisation of the task.

**Keywords:** Interaction, Virtual Reality, Degrees of Freedom, User Study

**Index Terms:** H.5.1 [INFORMATION INTERFACES AND PRESENTATION]: Artificial, augmented, and virtual realities— [H.5.2]: INFORMATION INTERFACES AND PRESENTATION—User Interfaces

## 1 INTRODUCTION

Virtual Reality (VR) provides new interactive experiences by immersing users in 3-D environments thanks to stereoscopic displays. These new technologies spread out from the academic field and are now used in entertainment (e.g. in video games or 3-D films) or in scientific visualisation to make complex phenomenon more understandable (e.g. fluid behaviour or construction techniques). Stereoscopic displays can be combined with the use of six DoF devices (such as data gloves). In such a case, users are able to interact with their surrounding environment through natural and intuitive gestures.

Such 3-D interaction provides a direct manipulation of the 3-D objects. At first, such techniques were described as more intuitive and as precise as the use of the traditional keyboard / mouse [17]. However, recent studies seem to indicate that 3-D interaction does not provide the needed precision even for simple tasks such as pointing [8]. Possible explanations to this phenomenon may be the following:

1. The absence of physical support (e.g. the desk if using a mouse) induced by the use of 3-D interaction techniques necessitates constant mid-air manipulation. After several minutes, fatigue causes significant hands' tremor and significantly reduces users' precision;
2. In VR users have to manipulate an increasing number of DoF simultaneously. Although they interact in the 3-D space as in the real world, the tasks they are required to perform in VR are slightly different from what they are used to. Contrary to VR, in everyday interactions, the physical constraints inherent to our world reduce the dimensionality of the tasks (i.e. the number of DoF involved in the task). For instance to position an object in VR users manipulate three DoF, whereas in the real world gravity makes an object lay on a support thus reducing the dimensionality of the task to two dimensions.

A recent trend in the field of Human - Computer Interaction investigates the use of TSI for 3-D interaction to try to overcome these issues [11] [15]. Interaction techniques based on TSI introduce physical constraints that reduce the dimension of the manipulation space and provide a physical support during the manipulation. Thus, it can provide an answer to both problems identified above. Several studies have already proved that the use of interaction techniques based on TSI can lead to satisfactory performance [15]. However, they were not interested in understanding how these techniques influenced the task. For the purpose of providing the users with the best 3-D interaction experience, we think it is important to better understand this phenomenon. In the present paper, we provide empirical evidences of the influence of the use of TSI for 3-D interaction. This paper focuses on the influence of the decomposition and of the physical support on the performance.

The remainder of this paper is organised as follow. In the next section, we present the relevant literature related to our purpose. Then, in Section 3, we describe the experimental study we have conducted to investigate the relative impact of the physical support and the task's decomposition induced by the use of TSI on the users' performance. In Section 4 and 5, we provide the results we obtained and their analysis. Finally, in Section 6 we conclude this work.

## 2 RELATED WORKS

In this section we present the relevant works investigating the impact of the configuration of DoF on the users' performance. We also present the solutions which have already been proposed to try to improve the control of users during 3-D positioning, including TSI.

### 2.1 Degrees of freedom

To describe the impact of the manipulation of multiple DoF during an interaction task, Jacob et al. [13] introduce the notion of *perceptual structure*. Each task has its own perceptual structure that gathers all the attributes that can be modified. It can be *integral* if the semantic distance between the attributes is low or *separable* if the semantic distance is important.

Jacob et al. studied the influence of the perceptual structure on the performance users reach. They showed that to obtain the best performance the device should implement a behaviour that matches

*e-mail:mveit@unistra.fr
†e-mail:a.capobianco@unistra.fr
‡e-mail:bechmann@unistra.fr

the perceptual structure of the task. However, this framework ignores the users possible inability to manipulate all the DoF simultaneously if their number increases. It has been shown that if the task involves the manipulation of many DoF at the same time, such as during a docking task, users decompose the task into several, less complex subtasks, involving fewer DoF [19].

Casiez et al. [7] compared two elastic devices, the first one (the Spacemouse ©) integrating and the second one (the DigiHaptic) separating the manipulation of the three DoF involved in a manipulation task. They showed that it was strenuous for the users to coordinate the manipulation of two and three DoF during a steering task even using the SpaceMouse. Moreover, if the steering task required users to manipulate the three DoF at the same time, the users reached the same level of coordination using both devices. This suggests that the users try to control the DoF separately even if the perceptual structure of the task is integral. Berard et al. [5] confirmed this idea in an experimental study. They compared four devices in a 3-D positioning task: the mouse, the DepthSlider, the SpaceNavigator and the Wii Remote™tracked by a Vicon™. The obtained results show that the mouse and the DepthSlider are more efficient than 3-D input devices for precise 3-D placement. However, they do not investigate the impact on the performance of the characteristics of the DoF decomposition (e.g. does separating the height and separating the depth lead to similar performance ?).

In this paper we are interested in the impact of the strategy of separation of the DoF on the users' performance. In particular, we focus on the impact of the separation of the depth dimension.

### 2.2 3-D positioning task and TSI

Several proposed techniques are based on the *Two-Component Model* [20]. This model states that the reaching movement is decomposed in two phases: the *ballistic phase* and the *control phase*. The ballistic phase corresponds to quick movements toward the target. The aim is to get coarsely near the target position. The control phase consists of several small corrective movements. During this phase, the user performs successive adjustments to get progressively close to the target. The **PRISM** technique, proposed by Frees et al. [10], is a non-linear filtering method implementing the Two-Component Model. It consists in modifying the C:D ratio according to the current speed. This method proved to be more precise than the direct manipulation. However, PRISM does not consider the possible strategies based on DoF decomposition used to complete the task. Such strategies may be implemented by the users to handle specific difficulties linked to 3-D interaction.

When interacting in a Virtual Environment (VE) the *depth* dimension presents such specific difficulties. Dellisanti et al. [8] highlight the anisotropic nature of the pointing task in VE. In their work, they show that it is harder for users to position the object along the depth than along the two other dimensions. They propose to adapt the **snap-to-grid** technique (the position of the pointer is rounded to the nearest point on a 3-D virtual grid) to try to counter balance the anisotropy and provide the users with more precision. However, this technique limits the position of the objects to a definite set of positions in the environment.

Recent works propose in order to decompose the task to improve the users' performance. The **Balloon Selection** technique proposed by Benko et al. [4] is based on the decomposition of the positioning task. Using this technique, the positioning of the selection cursor is decomposed into two subtasks: move the cursor into the horizontal plane and modify the height of the cursor. Benko et al. were surprised by the performance offered by their technique compared to the direct manipulation. In addition to decreasing the error rate, their technique does not increase the achievement time although it compels users to decompose the task. However, they do not propose any explanation for their results. Since it is based on TSI, their technique introduces both a decomposition and a physical support, but their study does not allow them to analyse the respective influence of these two factors on the users' performance.

Other works propose new interaction techniques based on TSI and providing a manipulation which is as direct as possible [14]. These works usually try to exploit the use of multiple contact points to allow the manipulation of all the DoF simultaneously. This is the case of the **Shallow-Depth Interaction** [11]. Using this technique, the more the contact points, the more the number of DoF are manipulated at the same time. In this work, Hancock et al. compared the one, two and three-touch input techniques during a passing and docking task. The results show that the three-touch technique leads to smaller achievement times and is less error prone than the two and one-touch techniques. However, they are not interested in the possible strategies based on a progressive decomposition of the task using the three-touch technique.

Martinet et al. [15] proposed two techniques for 3-D positioning: the **Z-technique** and the **Multi-Touch Viewports** technique. The Z-technique consists of manipulating the object in the X-Y plane using the dominant hand and separating the Z axis using the non-dominant hand. The other technique is based on the manipulation into four different projected views. They show that both techniques lead to similar achievement times and precisions during 3-D positioning. However, they did not study the impact of the way the task is decomposed on the users' performance.

In the next section, we present the experimental study we have conducted, in order to analyse the respective influence of both the presence of a physical support and the decomposition of the task's DoF induced by the use of a TSI.

## 3 Experimental study

In this experiment, we want to study two factors: the presence of the physical support and the type of decomposition. These factors may have specific influences on the performance of the users. On the one hand, the physical support may provide passive haptic cues (i.e. the haptic information is not controlled and depends only on the physical properties of the interaction device) which could complement visual information for a better evaluation and realisation of the task. On the other hand, the separation of the DoF may be appropriate to the perceptual structure of the task, influencing the users' performance according to the results obtained by Jacob et al. [13].

A full study would have necessitated to compare three (1-D, 2-D, 3-D positioning tasks) $\times$ three (integrated manipulation, depth axis separation and height axis separation) conditions. In order to limit the number of conditions we chose to focus on 3-D tasks. In that context, it is impossible to introduce a TSI without introducing any decomposition of the task. This is only possible for 2-D input tasks. In order to fully validate our results, it might be necessary to complete the design of this first study in a further experiment. In particular, a study on 1-D and 2-D tasks is needed to fully separate passive haptic feedback and decomposition. In this study, we can not completely isolate the physical support factor.

### 3.1 Interaction techniques

To investigate the impact of the physical support and the type of decomposition, we compare the following five techniques. None of them introduce a collocated manipulation (i.e. that the position of the device and the position of the manipulated object are not confounded).

#### 3.1.1 Direct manipulation (DM)

Using this technique, the user pinches the index and the thumb to define the position of reference. By moving the hand while maintaining the two fingers pinched, the user moves the object. Using this technique, we introduce no physical support and no decomposition.

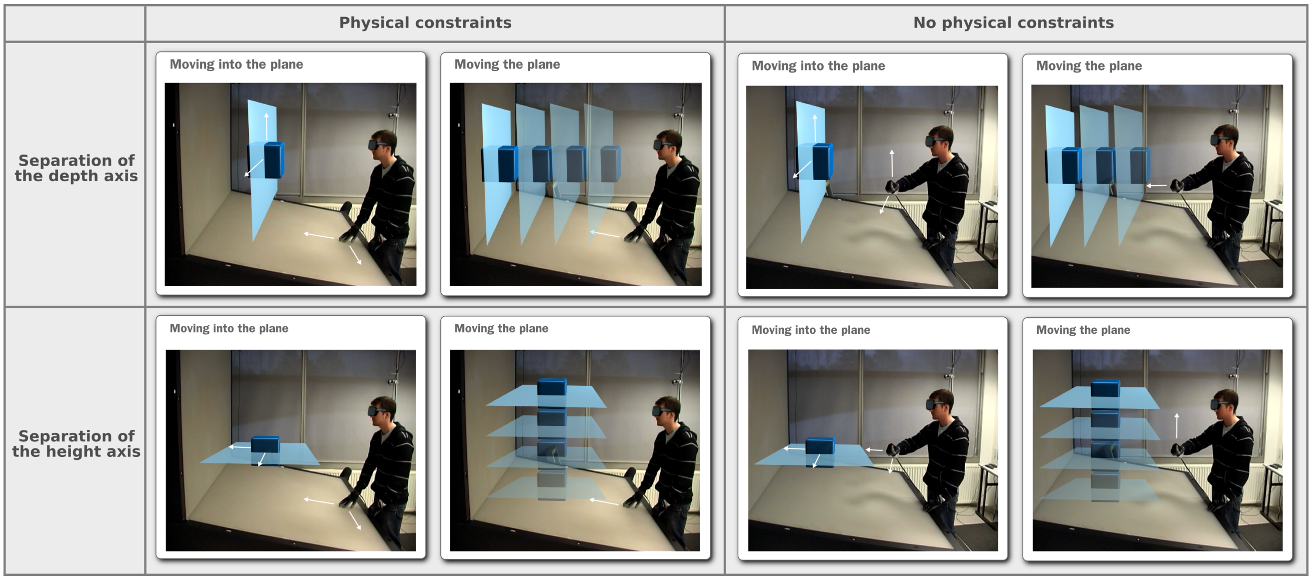


Figure 1: The four interaction techniques used to investigate the relative impact of the decomposition and the physical constraints.

#### 3.1.2 Direct manipulation separating the depth axis (DD)

We force users to decompose the task. By pinching the index and the thumb, the user can move the object in the X-Y plane (also called the manipulation plane). To move the object along the depth axis, the user pinches the index and the major. If the user tries to pinch the three fingers together, none of the movements carried out are applied to the object. The technique is illustrated on the top right part of the Figure 1. This technique separates the depth dimension but introduces no physical support.

#### 3.1.3 Direct manipulation separating the height axis (DH)

This technique is similar to the technique described in Section 3.1.2. The only difference is that it isolates the manipulation of the height dimension and the manipulation plane is the plane X-Z. The technique is illustrated on the bottom right part of the Figure 1. This technique separates the height and introduces no physical support.

#### 3.1.4 Touch screen separating the depth axis (TSD)

By asking the user to touch the screen at arms' reach, we introduce a physical support in addition to the decomposition. To move the object in the plane X-Y, the user slips his hand while touching the screen with the index finger. By touching the screen with the major finger, the user moves the object according to the depth axis. The technique is illustrated on the top left part of the Figure 1. This technique introduces both a separation of the depth dimension and a physical support.

#### 3.1.5 Touch screen separating the height axis (TSH)

This technique is similar to the one described in Section 3.1.4. Using this technique, the manipulation plane is the X-Z plane and the separated dimension is the height. The technique is illustrated on the top right part of the Figure 1. This technique introduces both a separation of the height axis and a physical support.

### 3.2 Tasks

We asked the subjects to achieve a *precise positioning task*. It consists in positioning a manipulated object, represented here by a cross aligned on the axes of the virtual world frame of reference, in a target, represented by a sphere of fixed radius and position. The radius of the sphere is 7.5 cm. The length of the axes is equal to the double of the radius of the sphere (i.e. 15 cm). In order to help the subjects to reach high precisions, the environment provides several depth cues (perspective, occlusion, head tracking and stereoscopy) in all conditions [6]. Moreover, we added visual cues to the manipulated object by highlighting the parts of the cross inside the sphere by a change in its colour. This allows the subjects to have a good perception of the relative position of the cross and the target. The lack of depth cues during the ballistic phase might not have influenced the results since the two objects were positioned at the same place for each new positioning. An illustration of both the cross and the sphere is given in Figure 2.

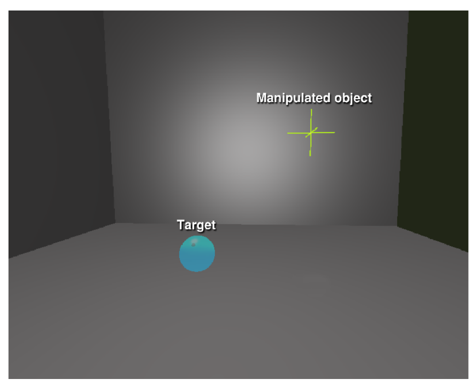


Figure 2: The visual feedback associated to the precise positioning task. The cross represents the manipulated object. The centre of the sphere represents the target the user has to reach.

The initial position of the cross is defined so that:

1. It is behind the target;
2. The segment between the initial position and the target is a perfect diagonal (i.e. the initial distance according to each dimension are the same). This was chosen to make sure that the 3 DoF where equally involved in the realisation of the task.

Each subject carried out six positioning for each of the four diagonals. The initial distance between the object and the target is 90 cm.

### 3.3 Experimental procedure

Fifteen voluntary subjects (students and peoples working at the University) took part in this study. They have various levels of experience in VR (ranging from subjects that discovered the environment during the experiment to subjects that occasionally use the environment). Each subject carried out twenty four positioning with the five techniques. Initially, the environment is presented to the subject. Then, a training session begins. This session consists in five blocks of five positioning, one block per technique. Before each block, the interaction technique is presented orally to the subject. During the training, no limit of time is imposed and we asked the subjects to take their time to understand how each interaction technique works. The objective is to familiarise the subject both with the environment and the various techniques. During the study, the training session lasted about 20 minutes for each subject.

In this study, we used a within subject design. Each subject achieved the twenty four positioning with the five techniques. To avoid any ordering effect, we used a Latin square distribution of the order of presentation of the techniques.

### 3.4 Measurements

During this experiment, we took the following measurements.

- Achievement time (AT, in second): time elapsed between the first and the last movement of the cross;
- Euclidean error (E, in mm): the euclidean distance between the manipulated object and the target at the end of the task;
- Error along X, Y and Z dimensions ($E_x$, $E_y$, $E_z$, in mm): error according to each dimension at the end of the task;
- Coordination: we used MDS (Magnitude of DoF Separation) [18] to compute NDC (Number of DoF Combined) a simple mathematical transformation of the measurement. In this case, $NDC = 1 + (2 \times (1 - MDS))$. It provides a real value (varying between 1 and 3) quantifying the number of DoF combined for each displacement of the manipulated object;
- Percentage of time spent to handle 1, 2 and 3 DoF: we use NDC to compute the proportion of time the user manipulates one, two and three DoF. If NDC $\in [1;1.66[$ the user manipulates one DoF. If NDC $\in [1.66;2.33[$ the user manipulates two DoF. If NDC $\in [2.33;3]$ the user manipulates three DoF;
- Percentage of time spent to handle 1, 2 and 3 DoF before and after a precision threshold: the threshold is defined as $5 \times E$. We use this threshold to take into account the precision of the various users. By differentiating the coordination before (i.e. far from the target) and after (i.e. near the target) the threshold, we want to observe if there is any specific strategies for each phase of the positioning regarding the DoF coordination.

At the end of each experimentation, the subjects were invited to fill in the NASA-TLX questionnaire [12] and to evaluate their preference regarding each technique on a 7-Likert scale (1 = I strongly disliked this technique ; 7 = I really appreciated this technique).

### 3.5 Hypotheses

We had the following expectations about the influence of the physical support and the decomposition on the users' performance.

**H1**: The physical support increases the precision and reduces the achievement time. The physical support may remove parasitical movements by reducing the fatigue and provides passive haptic information to the user.

**H2**: The integration decreases the precision but reduces the achievement times compared to the decomposition of the task. By combining the DoF, the user is able to benefit from an induced parallelism.

**H3**: By allowing the user to focus on a single dimension of manipulation, the separation of the DoF induces an increased precision linked to an improved motor control of the gesture according to the isolated dimension.

### 3.6 Apparatus

Each experiment was conducted on a Barco Consul, a VR environment providing stereoscopic display. The two screens have a resolution of $1400 \times 1050$ with a refresh rate of 120 Hz. The CrystalEyes CE-2 glasses are synchronised with the screens to ensure the stereoscopy. The images were generated using two workstations (one for each screen) fit out with a 3 GHz Intel Quad Core and a NVIDIA Quadro FX 5600. The head and the hands tracking were provided by an ART tracking system and optical trackers. The users used two XIST Data Gloves to interact with the VE.

## 4 Results

We analysed the data using an ANOVA with repeated measurements with one factor, the interaction technique (IT). When IT seems to have an influence on the dependent variable, we use a pairwise comparison with a Benjamini and Hochberg correction [3]. This correction controls the false discovery rate, the expected proportion of false discoveries amongst the rejected hypotheses. The false discovery rate is a less stringent condition than the family wise error rate, so this method is more powerful than others correction.

### 4.1 Performances

The performance are indicated by the measured precisions and achievement times.

#### 4.1.1 Precision

| IT | E | $E_x$ | $E_y$ | $E_z$ |
|---|---|---|---|---|
| *DM* | 6.94 mm | 1.64 mm | 2.52 mm | 5.29 mm |
| *DH* | 6.92 mm | 1.63 mm | 2.64 mm | 5.14 mm |
| *TSH* | 6.77 mm | 1.41 mm | 2.40 mm | 5.37 mm |
| *DD* | 5.95 mm | 1.48 mm | 2.27 mm | 4.35 mm |
| *TSD* | 5.93 mm | 1.43 mm | 2.19 mm | 4.45 mm |
| *F* | $F(4,51) = 4.6$ | $F(4,51) = 2.2$ | $F(4,51) = 2.9$ | $F(4,51) = 3.7$ |
| p | 0.003 | 0.079 | 0.029 | 0.048 |

Table 1: Mean errors.

Table 1 indicates the mean error for each technique and the average error according to each of the three dimensions. The interaction technique seems to have an effect on E. It seems that there is a significant difference between DM and the techniques separating the depth (DM vs. DD: p<0.001; DM vs. TSD: p<0.001). DD and TSD are respectively 14.26 % and 14.55 % more precise than DM. However, there is no difference between the techniques separating the height and DM (DM vs. DH: p=0.94; DM. vs. TSH: p=0.71). There is also a significant difference between the techniques separating the depth and those separating the height (DH vs. DD: p <0.001; DH vs. TSD: p<0.001; TSH vs. DH: p=0.0017; TSH vs. TSD: p=0.0016). The techniques DD and TSD are in average 13.22 % more precise than the DH and TSH techniques. These results suggest that the decomposition has an influence on the users' precision.

Indeed, for the techniques separating the depth the precision regarding the isolated dimension is significantly increased (DD vs. DM: p=0.003; DD vs. DH: p=0.008; DD vs. TSH: p=0.002; TSD vs. DM: p=0.005; TSD vs. DH: p=0.02; TSD vs. TSH: p=0.003). This is not the case for the techniques separating the height. Furthermore, we do not observe any difference between DH and TSH or between DD and TSD (DH vs. TSH: p=0.71; DD vs. TSD: p=0.94). These results suggest that the increased performance observed for DD and TSD cannot be explained by an improved motor control of a single dimension induced by the separated manipulation of the isolated DoF. They also suggest that there is no influence of the physical support on the users' precision in the context of a task's DoF separation.

#### 4.1.2 Achievement time

| IT | AT |
|---|---|
| *DM* | 11.24 s |
| *DH* | 16.88 s |
| *TSH* | 16.89 s |
| *DD* | 15.19 s |
| *TSD* | 15.36 s |
| *F* | $F(4,51) = 24.8$ |
| p | <0.001 |

Table 2: Mean achievement time.

There is a statistically significant influence of IT on AT (see Table 2). We observe an increase of almost 50 % for DH and TSH and an increase of 35.9 % for DD and TSD. We can also notice that DD and TSD are slightly quicker than DH and TSH. All the differences are statistically significant (p<0.001) except the differences between DH and TSH (p=0.97) and between DD and TSD (p=0.76). The type of decomposition seems to have an influence over AT. However, this is not the case for the physical support.

### 4.2 Behavioural analysis

The measurements about the coordination enable us to analyse positioning strategies employed by the subjects.

#### 4.2.1 Coordination

Table 3 indicates the level of coordination for each technique. The bigger the proportion of one DoF the more the users decompose the task. The interaction technique has an influence on all of these measurements. Except for the differences between DH and TSH and between DD and TSD, the differences are significant for the percentages of times spent to handle one and two DoF (p<0.001). Particularly, we notice that the percentages of time spent to handle one and two DoF for the techniques separating the height are more important than for those separating the depth (p<0.001). This may suggest that users try to separate the depth in the manipulation plane for DH and TSH.

With regard to the percentage of time spent to handle three DoF, as the percentage of time is null for the four techniques decomposing the task, only the differences between DM and the other techniques are significant (p<0.001).

| IT | 1 DoF | 2 DoF | 3 DoF |
|---|---|---|---|
| DM | 12.39 % | 51.89 % | 34.39 % |
| DH | 76.82 % | 22.94 % | 0.0 % |
| TSH | 77.88 % | 22.20 % | 0.0 % |
| DD | 69.12 % | 30.72 % | 0.0 % |
| TSD | 69.68 % | 30.35 % | 0.0 % |
| *F* | $F(4,51) = 675.8$ | $F(4,51) = 100.3$ | $F(4,51) = 346.5$ |
| p | <0.001 | <0.001 | <0.001 |

Table 3: Mean proportion of time spent to handle one, two and three DoF in relation to the manipulation time.

#### 4.2.2 Decomposition of the task

Based on the Two-Component Model, we want to study the behaviour of the users regarding the decomposition both in the ballistic and the control phase. To this purpose, we distinguish between the positioning near and far the target. Table 4 indicates the mean percentage of time spent to handle one, two and three DoF during these two phases. For each technique, we compared the percentage of time before and after the threshold using a t-test in order to figure out if the observed difference is significant. For all the comparisons (except the comparisons of the percentages of time spent to handle three DoF for the techniques decomposing the task) the values of p are lower than 0.001. We can notice that once the threshold is passed, the percentages of time spent to handle one DoF at a time increases and the percentages of time spent to handle two and three DoF decreases. This suggests that the users try to decompose the positioning even with DM.

| IT | 1 DoF | | 2 DoF | | 3 DoF | |
|---|---|---|---|---|---|---|
| | Before | After | Before | After | Before | After |
| DM | 11.57 % | 15.68 % | 47.30 % | 56.89 % | 41.12 % | 27.43 % |
| DH | 75.80 % | 78.69 % | 23.99 % | 21.22 % | 0.0 % | 0.0 % |
| TSH | 76.22 % | 80.12 % | 23.73 % | 19.85 % | 0.0 % | 0.0 % |
| DD | 65.22 % | 81.64 % | 34.61 % | 18.30 % | 0.0 % | 0.0 % |
| TSD | 65.95 % | 81.97 % | 34.01 % | 18.01 % | 0.0 % | 0.0 % |

Table 4: Mean proportion of time spent to handle one, two and three DoF before and after the precision threshold.

Moreover, we can notice the strong proportion of time spent to handle one DoF at the end of the task. For DH, TSH, DD and TSD, it appears that the users decompose the positioning in the manipulation plane. In order to study the strategy used for the positioning in the manipulation plane, we compute the proportion of time spent to handle one and two DoF when moving the object in the manipulation plane.

| IT | 1 DoF | | 2 DoF | |
|---|---|---|---|---|
| | Before | After | Before | After |
| *DH* | 45.74 % | 52.26 % | 54.26 % | 47.73 % |
| *TSH* | 47.81 % | 55.33 % | 52.18 % | 44.67 % |
| *DD* | 26.88 % | 41.99 % | 73.12 % | 58.01 % |
| *TSD* | 28.32 % | 42.36 % | 71.68 % | 57.63 % |
| *F* | $F(3,38) = 14.5$ | $F(3,38) = 7.2$ | $F(3,38) = 14.5$ | $F(3,38) = 7.2$ |
| p | <0.001 | <0.001 | <0.001 | <0.001 |

Table 5: Mean proportion of time spent to manipulate one and two DoF in the manipulation plane before and after the threshold.

The interaction technique has an influence on the decomposition in the manipulation plane. Table 5 shows that the integration in the manipulation plane is more important for the techniques separating the depth than for the techniques separating the height. The results also show that the decomposition in the manipulation plane is more important at the end of the task (p<0.001). This result suggests that even with the techniques decomposing the task, the user separates the 2-D positioning in the manipulation plane at the end of the task.

### 4.3 Subjective evaluation

For the NASA-TLX questionnaire, we extract the evaluation for each criterion in addition to the subjective evaluation of the users' preference.

#### 4.3.1 Preference

Table 6 indicates the average evaluation for each technique. Only the differences between DD and TSD and between DH and TSH are not significant. The other differences are significant (p<0.05). We can notice that the evaluation of DM is better than those of DH and TSH. In opposition, DD and TSD have a better evaluation than

| IT | USE |
|---|---|
| *DM* | 4.80 |
| *DH* | 3.73 |
| *TSH* | 3.67 |
| *DD* | 5.67 |
| *TSD* | 5.80 |
| *F* | $F(4,56) = 13.7$ |
| p | <0.001 |

Table 6: Mean USE (User Subjective Evaluation) mark.

DM (DM vs. DD: p=0.023; DM vs. TSD: p=0.01). It appears that the separation of the depth is appreciated by the users.

#### 4.3.2 Workload

NASA-TLX questionnaire provides an evaluation according to various criteria. By combining these factors, it is possible to compute the Weighted Workload (WWL) which estimates the cognitive load. The results regarding the various criteria and WWL are given in Table 7.

The physical support seems to reduce the fatigue (concerning the physical demand, DH vs. TSH: p=0.02; DD vs. TSD: p=0.18). However, it seems that it has no influence on any of the other factors (p>0.1). The decomposition seems to have an influence on WWL. The decomposition of the task in an X-Z positioning and a Y positioning makes the task more complex. This is not the case for the X-Y plus Z decomposition which does not seem to increase WWL. Except the difference between DD and DM (p=0.078), the differences between the techniques separating the depth and the other techniques are significant with the values of p lower than 0.05. Moreover, the decomposition seems to have a significant influence on the level of frustration (the comparisons between the techniques separating the height and the depth are significant, p<0.001).

## 5 Discussions

We discuss in this section the influence of both the physical support and the decomposition introduced by TSI.

### 5.1 Influence of the physical support

Among the assumptions of this work, we thought that the introduction of a physical support could help the users to have a better control of their gestures. In particular, we thought that it could have an influence on the performance regarding precision by helping the users to remove parasitical movements due to possible tremors of the hand. It also introduces passive haptic feedback during the achievement of the task. We thought that this information, enriching the visual information, could have helped the users to correct the position of the object during the control phase. However, our results show that the presence of a physical support has no influence on the precision. If the precision differs between the techniques separating the depth and the height, the in-between comparison assessing the influence of the presence of the physical support showed no statistical evidence (see Table 1). We think that this can be explained by the fact that subjects mainly rely on visual information to evaluate the precision of the positioning of the object. Meyer et al. [16] conducted a study that indicates that during the control phase, the visual information is predominant. Likewise, it seems that the haptic cues were not exploited by the subjects to achieve more precision.

Subjective evaluation marks underlines that the introduction of a physical support usually improves the evaluation of the subjects. However, the only statistically significant effect we can notice is the reduction of the fatigue. We think that it is possible, with more numerous subjects and longer manipulation sessions, that the differences in terms of subjective preferences may have been more strongly marked, especially concerning performance (P), effort (E) and WWL.

We did not find any evidence of an influence of the physical support regarding achievement times. Techniques separating either height or depth lead to similar results with or without the physical support regarding achievement times (see Section 4.1.2). We can then conclude that we were wrong about our first hypothesis: contrary to our expectations, the presence of a physical support does not lead to increased performance during positioning tasks. It only contributes to strongly reduce the fatigue the users might experiment during long sessions of manipulation. At this point, it is important to underline that our experimental protocol does not allow finger-tip precision. The manipulation is mainly carried out by the user's shoulder, elbow and wrist, i.e. the gross joints and muscle groups of the limb. The smaller, finer joints and muscle groups on the fingers were not utilised for the manipulation. Zhai et al. [21] showed that implication of smaller groups leads to increased performance during positioning tasks and can allow pixel level pointing precision [1]. Thus, using a device allowing finger-tip manipulation may have led to increased precision. Moreover, it is likely that in the context of a 1-D or a 2-D task, without the separation, the subjects strategies might have been different, leading to different results.

In the context of 3-D tasks, it seems that TSI would be adapted for long sessions of work. The reduction regarding the physical demand could have been more important as the duration of the test would have increased. The benefits measured here are for one hour sessions of work. This remains relatively short compared to the length of the sessions of work in the context of geometric modelling applications for instance.

### 5.2 Influence of the decomposition

It appears that the decomposition is the factor having the most important influence on the results. It impacts achievement times, precision and subjective evaluations. We will discuss each of these aspects here.

#### 5.2.1 Achievement times

Not surprisingly, we found that the techniques introducing a decomposition induce more important AT than the technique integrating the manipulation of the DoF. By making it possible to handle the three dimensions of the space simultaneously, DM allows to follow a direct trajectory toward the target. This result seems to indicate that the DM technique is appropriate during the ballistic phase of the positioning task. These results are exaggerated since the task consists of a true 3-D positioning which constitutes the worst case for the techniques decomposing the task. The differences observed are thus partially due to the fact that a separation of the DoF introduces a minimal distance to the target that is more important. By facilitating the covering of long distances (thanks to a non-linear transfer function such as PRISM [10]) this phenomenon could have been reduced.

A second factor also helps to explain this difference. AT is defined as the time between the first and the last displacement of the object. However, during the experiments we notice that the users spend more time analysing the task with DD, TSD, DH and TSH than with DM. This phenomenon appears at the end of the task when they enter the control phase. The users spent a significant time to identifying the dimensions to be corrected to reach the maximum of precision. It seems that this phenomenon appears to a lesser extent with DM. This is confirmed by the subjective evaluation regarding the performance (the P factor in the NASA-TLX questionnaire). The users evaluated the techniques separating the depth as providing more precision than DM.

We can also observe a difference between the techniques separating the depth and the techniques separating the height. With the

| IT | MD | PD | TD | P | E | F | WWL |
|---|---|---|---|---|---|---|---|
| *DM* | 2.40 | 5.13 | 3.53 | 4.27 | 3.67 | 3.40 | 3.77 |
| *DH* | 4.67 | 4.67 | 3.73 | 4.20 | 4.80 | 5.27 | 4.83 |
| *TSH* | 4.33 | 2.80 | 3.00 | 3.40 | 4.00 | 4.93 | 4.02 |
| *DD* | 2.00 | 3.80 | 2.73 | 2.67 | 2.73 | 2.20 | 2.67 |
| *TSD* | 2.33 | 2.67 | 2.27 | 2.07 | 2.53 | 2.13 | 2.31 |
| $F$ | $F(4,56) = 9.8$ | $F(4,56) = 5.8$ | $F(4,56) = 2.7$ | $F(4,56) = 5.2$ | $F(4,56) = 5.2$ | $F(4,56) = 9.5$ | $F(4,56) = 7.5$ |
| p | <0.001 | <0.001 | 0.03 | <0.001 | <0.001 | <0.001 | <0.001 |

Table 7: *Mental Demand* (MD), *Physical Demand* (PD), *Temporal Demand* (TD), *Performance* (P), *Effort* (E) and *Frustration Level* (FL).

techniques separating the depth, the users are able to achieve the task on average 1.61 seconds quicker than with the techniques separating the height. For the techniques DH and TSH, the positioning in the horizontal plane does not allow to separate the depth. Indeed, a depth displacement generally implies a lateral displacement and conversely. Therefore, depth adjustments are more difficult to carry out with DH and TSH.

These results may indicate that the perceptual structure of the DD and TSD is more appropriate to the task than the perceptual structure of the DH and TSH techniques. If this is the case, we should observe similar results regarding the precision users reached.

### 5.2.2 Precision

Concerning precision, our data shows contradictory results between the techniques separating the depth axis and the techniques separating the height axis. Compared to DM, the DD and TSD techniques introduce a significant benefit regarding the precision, while DH and TSH seem to lead to similar performance. If we look at the results separately for each dimension (see Table 1), we notice that the increases of precision for the techniques separating the depth mainly relies on a reduced error regarding $E_z$. However, we do not observe any improvement for the techniques separating the height according to the Y dimension, which is in opposition with our third hypothesis.

This hypothesis relied on the assumption that the separation of the manipulation of specific dimensions would bring an increased control regarding that dimension. If that was the case, we should have observed benefits regarding the $E_y$ axis for the DH and TSH techniques. Therefore, the observed benefit may not come from a better motor control. Longer training or longer test sessions might have influenced these results. However, we think that the behavioural analysis we conducted may explain this phenomenon. Our results indicate that with DH and TSH, the subjects behaved as if they had to realise a $3 \times 1$-D task, leading to increased proportions of time when the subjects manipulated only 1 DoF. This is also accompanied by a strong increase of the mental demand for the DH and TSH techniques. On the other hand, with DD and TSD they were more able to integrate the manipulation of the 2 DoF in the manipulation plane. We think that these results are explained by the fact that the perceptual structure of the DD and TSD techniques is closer to the perceptual structure of the task than DH and TSH. This may be especially true at the end of the task, when the subjects are trying to be precise.

Previous studies on 3-D positioning tasks, such as the evaluation of the Rockin'Mouse, seem to indicate that the perceptual structure of the positioning task varies according to the phase of the realisation. Indeed, in their experimental study Balakrishnan et al. [2] found that the movements during the control phase are less coordinated than the movements during the ballistic phase. The results of Balakrishnan et al. and ours show that during the ballistic phase, the perceptual structure is integral. Handling the three DoF at the same time is appropriate to this phase which consists of approaching the target coarsely. However, during the control phase the user needs precision, the simultaneous handling of the three DoF does not seem suitable. The user does not conceive the 3-D positioning task as a whole any more. The user may want to perform successive adjustments according to each dimension (and particularly according to the depth axis). The perceptual structure in this case is separable and may be conceived as three distinct and independent sub-tasks. This assumption appears to be confirmed by the analysis of the subjective evaluation and the coordination of the DoF. However, we need to perform additional experimental studies to investigate these questions and identify the perceptual structure of the positioning during both phases. One of our hypothesis is that the semantic information associated with the DoF is not the only criteria used to conceptualise a task ; the difficulty of manipulation may be another. We also think that visual information may influence the users' strategies during 3-D positioning. Enriching the users' visual feedback, such as using semi-transparent intersection planes, may help users to evaluate specific dimensions and separate these dimensions thus influencing the perceptual structure.

There may be one concern regarding the orientation of the screen used for the TSI that may have an influence on the performance. The screen being almost horizontal, the TSD technique seems to offer a better *stimuli - response compatibility* [9] than the TSH technique. However, we do not observe any difference neither regarding the euclidean distance nor regarding the achievement times between DD and TSD or DH and TSH. This indicates that the orientation of the screen has a little influence on the performance. This is confirmed by the subjective evaluation since we do not observe any significant difference between DD and TSD or DH and TSH regarding the mental demand ($p>0.5$). Moreover, in the study conducted by Martinet et al. [15], they compared the Z-technique to the Multi-Touch Viewports technique. In the Z-Technique, the depth displacements are orthogonal to the users' hand movements, which is similar to the use of a vertical touch screen in our experiment. They showed that the two techniques performed similarly even if the Multi-Touch Viewports offers an optimal stimuli-response compatibility. Their result, like ours, suggest that orientation mismatch between the screen and the manipulation space may have a little influence on the performance, maybe because of the non-collocated manipulation. However, their study and ours do not take into account possible learning effects that might influence performance over long sessions.

### 5.2.3 Subjective evaluation

The results regarding the subjective evaluations seem to confirm some of the hypotheses proposed in the previous section. In general, the decomposition of the task was perceived by the users as a handicap in the ballistic phase. During this phase, they prefer the DM technique. However, during the control phase, the decomposition was perceived as an advantage. Many remarks were made about the fact that the separation of the depth seems much more natural than the separation of the height. This is confirmed by the analysis of the subjective preference of the users. The users allotted lower marks to the techniques separating the height. However, the subjective evaluation shows a clear preference for TSD and DD, including regarding DM. That seems to confirm the appropriateness of the depth separation.

Initially, we thought that the decomposition would increase WWL, in particular by introducing a more complex phase of planning. However, it seems that it depends on the decomposition used. Once again, the separation of the depth does not seem to introduce any increase concerning WWL compared to DM. This result may show the appropriateness of the perceptual structure of the technique to the task. The techniques separating the height seems to increase WWL significantly. We think that this result is due to the lack of intuitiveness of the height separation and its inappropriateness to the difficulties of the task. This results in a very high level of frustration, which is of 5.1 in mean. The depth being the most complex dimension to apprehend in a 3-D environment, the users seem to benefit from its separation for the achievement of the task.

### 5.3 Touch screen interaction and 3-D positioning

The results we obtained indicate that touch screen interaction may be beneficial in the context of 3-D positioning tasks.

- If users need to perform precise positioning, while it necessitates the introduction of a decomposition, the use of touch screen interaction seems to be efficient;
- If users may perform long sessions of manipulation, the use of touch screen interaction significantly reduces the amount of fatigue.

However, the subjective evaluations suggest that these techniques are not appropriated for the ballistic phase. It is possible, using mobile devices for instance, to combine touch screen interaction with direct manipulation techniques in order to introduce physical constraints and decomposition during specific phases of the achievement of the task.

## 6 Conclusion

In the present paper, we investigated the use of TSI for 3-D interaction. We analyse the impact of the use of a physical support and the decomposition of a 3-D task induced by the use of TSI in the context of 3-D interaction. The results we obtained suggest that if the physical support does not influence the users' precision, the decomposition of the task has a strong influence. The interaction techniques separating the depth dimension lead to better precision than that of the others. Therefore, if the use of TSI for 3-D positioning imposes a decomposition of the task, it should separate the depth axis.

Our results also provide empirical evidences that the perceptual structure of the 3-D positioning task varies along the task realisation. It starts with an integral structure during the ballistic phase. It ends with a separable structure during the control phase. It also indicates that the semantic of the DoF is not the only criterion used to conceptualise a task: the difficulty of manipulation may be another.

## References


[1] P.-A. Albinsson and S. Zhai. High precision touch screen interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '03, pages 105–112, New York, NY, USA, 2003. ACM.

[2] R. Balakrishnan, T. Baudel, G. Kurtenbach, and G. Fitzmaurice. The rockin'mouse: integral 3d manipulation on a plane. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '97, pages 311–318, New York, NY, USA, 1997. ACM.

[3] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Royal Statistical Society*, 57(1):289–300, 1995.

[4] H. Benko and S. Feiner. Balloon selection: A multi-finger technique for accurate low-fatigue 3d selection. In *Proceedings of the IEEE Symposium on 3D User Interfaces*, 3DUI '07, pages 79–86, March 2007.

[5] F. Bérard, J. Ip, M. Benovoy, D. El-Shimy, J. R. Blum, and J. R. Cooperstock. Did "minority report" get it wrong? superiority of the mouse over 3d input devices in a 3d placement task. In *Proceedings of the IFIP TC International Conference on Human-Computer Interaction*, INTERACT '09, pages 400–414, Berlin, Heidelberg, 2009. Springer.

[6] D. A. Bowman, S. Coquillart, B. Froehlich, M. Hirose, Y. Kitamura, K. Kiyokawa, and W. Stuerzlinger. 3d user interfaces: New directions and perspectives. *IEEE Computer Graphics and Applications*, 28(6):20–36, nov 2008.

[7] G. Casiez, P. Plenacoste, and C. Chaillou. Does dof separation on elastic devices improve user 3d steering task performance? In *Proceedings of the Asia Pacific conference on computer human interaction*, APCHI '04, pages 70–80, Berlin, Germany, 2004. Springer.

[8] M. Dellisanti, M. Fiorentino, G. Monno, and A. Uva. Enhanced 3d object snap for cad modelling on large stereo displays. *International Journal of Computer Applications in Technology*, 33(1):54–62, 2008.

[9] P. Fitts and C. Seeger. S-r compatibility: spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 46(3):199–210, Sep 1953.

[10] S. Frees and G. D. Kessler. Precise and rapid interaction through scaled manipulation in immersive virtual environments. In *Proceedings of the IEEE Conference on Virtual Reality*, VR '05, pages 99–106, Washington, DC, USA, 2005. IEEE Computer Society.

[11] M. Hancock, S. Carpendale, and A. Cockburn. Shallow-depth 3d interaction: design and evaluation of one-, two- and three-touch techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, pages 1147–1156, New York, NY, USA, 2007. ACM.

[12] S. G. Hart and L. E. Stavenland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati, editors, *Human Mental Workload*, chapter 7, pages 139–183. Elsevier, 1988.

[13] R. J. K. Jacob, L. E. Sibert, D. C. McFarlane, and J. M. Preston Mullen. Integrality and separability of input devices. *ACM Transaction on Computer-Human Interaction*, 1(1):3–26, 1994.

[14] R. Kruger, S. Carpendale, S. D. Scott, and A. Tang. Fluid integration of rotation and translation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '05, pages 601–610, New York, NY, USA, 2005. ACM.

[15] A. Martinet, G. Casiez, and L. Grisoni. The design and evaluation of 3d positioning techniques for multi-touch displays. In *Proceedings of the IEEE Symposium on 3D User Interfaces*, 3DUI '10, pages 115 –118, march 2010.

[16] D. E. Meyer, R. A. Abrams, S. Kornblum, C. E. Wright, and J. E. K. Smith. Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological review*, 95(3):340–370, 1988.

[17] C. Michael, S. J. Mountford, and A. Sellen. A study in interactive 3-d rotation using 2-d control devices. *SIGGRAPH Computer Graphics*, 22(4):121–129, 1988.

[18] M. Veit, A. Capobianco, and D. Bechmann. Influence of degrees of freedom's manipulation on performances during orientation tasks in virtual reality environments. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '09, pages 51–58, New York, NY, USA, 2009. ACM.

[19] Y. Wang, C. L. MacKenzie, V. A. Summers, and K. S. Booth. The structure of object transportation and orientation in human-computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '98, pages 312–319, New York, NY, USA, 1998. ACM.

[20] R. Woodworth. The accuracy of voluntary movement. *Psychological Review*, 3:1–106, 1899.

[21] S. Zhai, P. Milgram, and W. Buxton. The influence of muscle groups on performance of multiple degree-of-freedom input. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '96, pages 308–315, New York, NY, USA, 1996. ACM.