Design of space-centered interaction using invisible and intangible spatial inputs

David Jean, Kwangjin Hong, Keechul Jung

(*Human-computer Interaction Lab, Graduate school of Soongsil university, Seoul 156-743, Korea*)

Abstract: In this paper, we investigate methodologies to improve direct-touch interaction on invisible and intangible spatial input. We firstly discuss about the motive of looking for a new input method for whole body interaction and how it can be meaningful. We also describe the role that can play spatial interaction to improve the freedom of interaction for a user. We propose a method of spatial centered interaction using invisible and intangible spatial inputs. However, given their lack of tactile feedback and visual representation, direct touch interaction on such input can be confused. In order to make a step toward understanding causes and solutions for such phenomena, we made 2 user experiments. In the first one, we test 5 setups of helper that provide information of the location of the input by constraining the dimension it is located at. The results show that using marker on the ground and a relationship with the height of the user's body improve significantly the locative task. In the second experiment, we create a dancing game using invisible and intangible spatial inputs and we stress the results obtained in the first experiment within this cognitively demanding context. Results show that the same setup of helper is still providing very good results in that context.

Key words: human-computer interaction (HCI); spatial centered interaction; whole body interaction; input method

CLD number: TP391.38 **Document code:** A
Article ID: 1674-8042(2012)02-0137-009 **doi:** 10.3969/j.issn.1674-8042.2012.02.0

0 Introduction

In recent human-computer interaction (HCI) area, new interaction models based on vision sensors have become an important research area. Thanks to the rapid evolution of computer's process power, fruitful researches in computer vision and mass availability of vision sensors, vision based interfaces are mature enough to produce interesting results and whole body based interaction is now considered as a concrete possibility. Popular research trends focus on the input acquisition based on postures and gestures. A posture is a static body position that can be matched to a template in order to be recognized. On the other hand, the gesture can be described as a succession of posture connected by motion over a short time span. From these two methods, we can categorize the posture-based methods as static interaction, because the process employed to reach the posture is not considered, only the results of the posture itself is. However, gesture recognition is a temporal based interaction, because all the movement is considered over time. As a matter of fact, a temporal interaction allows less freedom of use than a static interaction, because during the whole period of time the gesture is considered, any other activity can hardly be performed, as they could interfere with the gesture and provide an undesired result. In addition, a temporal interaction method allows a larger number of variations of input to be considered and more precision than that of its static counterpart, allowing for example to take into account mapping processes with more ease.

Although this precision is an important factor, there are some situations in which the freedom of use brought by a static based interaction must be considered as an even more important part of the system, especially in systems where explorability, a key usability component, plays an important role, like in entertainment or creative activities. This need is even more important since a system involves the whole body, where even a static based interaction like one based on posture, if it constrains the whole body, can still be highly limited. Interaction using vision sensors should provide an important advantage over other systems when dealing with user freedom of interaction, because it provides an unob-

※ **Received data:** 2011-10-12
Corresponding author: David Jean (Jeandavid31@gmail.com)

trusive way to interact without requiring any specific devices or clothes and because it should ideally be able to let the user perform any gesture in his environment. However, the interaction designs based on posture or gesture lead in fact to the other extreme and the user can actually feel totally constrained to have to perform a specific and precise gesture to successfully enter the desired input to the system.

Those issues arise because the user does not only generate the input anymore, but is himself the input. All the states that used to be due to the machine by a traditional device (specific position of the mouse, a certain key pressed...) are now due to the user. Therefore, all the constraints that lied with the input devices shift to lie with the user. But constraining directly the user takes away all his "freedom of interaction". "Freedom of interaction" describes the possibility that exists for a user to interact with a given input device in numerous ways he desires, motivated by any kind of purpose. Freedom of interaction plays an important role in many aspects of using a device.

Being able to use a device in numerous ways can firstly be seen as a matter of comfort of use, related to ergonomics. The possible use of the keyboard with one hand when the other is busy, not having to sit on a chair to quickly type a small message is situations that occur many times in everyday life. And although flexibility of a keyboard and a mouse can hardly be used as a reference, we can imagine that even more constraining devices could have impacted negatively on the adoption of computers. The keyboard-mouse duo also proves that freedom of interaction can be efficiently driven. Programming, writing any kind of text, navigating through a form can usually more effectively be done using only the keyboard with two hands. However, searching, selecting, copying and pasting a part of a text from a big document is usually faster using both mouse and keyboard at the same time.

Moreover, freedom of interaction is directly linked to the explorability factor for a given device, with direct consequences on creativity and entertainment. Once a certain level reached in an activity, it is common as a challenging process, to push the limits of the device, tools, interactors, instruments, etc. A good example can be found in music. Many have been amazed to see Jimi Hendrix plays its guitar behind his back, or using his tooth rather than his hands. The practice can appear non useful in a strict "sound making" point of view, and cannot be categorized as efficient nor comfortable way to play guitar. But by exploiting the explorability potential given by the high level of freedom of interaction of his instrument, he has been able to "push" further its limits. This is at the same time a way to produce new challenges, to discover a range of new possibilities, as well as generate admiration of people around. New challenges and possibilities offer an interesting space for creativity, and admiration of peers also plays an important role in social interaction that extend entertaining dynamic. Now, lots of guitar players (or players of video games involving a guitar shaped devices) are trying to miitate Jimi Hendrix, who is considered to be the ultimate challenge when mastering this instrument.

A parallel can be drawn from music performance to performance using new generation interface. Dance dance revolution is a game launched by Konamy in the late 90's which involves the user in a beat matching based game using mat pressure sensors, inviting the player to move its foot, or in other words, to "dance", in order to press these interactors. Its large success has created a popular community of players that pushes the limit of the game by creating their own new rules. The goal is not to take the most difficult song to get a perfect score, but to take an entertaining one and perform a challenging choreography involving all the body, far beyond the actual controllers on foot, which hence becomes only restrictions to respect. Moreover, a player using 2 arcade games or 2 players in cooperation can provide even more spectacular choreographies, switching from place to place or interacting with the two games at the same time.

Through these examples, we want to show the importance of freedom of interaction, that when is no respected, can lead to limitation of creation and involvement in the activity, such as in the case of body centered interaction.

In order to offer a new perspective and provide an alternative way to interact by using the whole body, while trying to still provide some freedom of interaction, we explore the challenge in applying and investigating an interactive method generated by a paradigm shift from a body-centered interaction to a space-centered interaction. Our objective is to alleviate the constraint lying on the user by displacing them to the space surrounding him. In order to do so, we propose an interaction method using invisible and intangible spatial inputs.

1 Invisible and intangible spatial inputs

An invisible and intangible spatial input(IISI) can be seen as a virtual input corresponding to a specific area in the space. This is a new type of input, however, from the author's knowledge, no previous studies have been done in order to investigate the possibilities of system based on such type of input in the context of a 3D whole body interaction as well as the new challenges generated by invisible and intangible properties. Such input is characterized by its position in the 3D real space, its shape, its volume, its behavior and its function. When something (any part of a user's body in our context of study) is detected inside the volume that defines the IISI, its associated function is triggered, similarly as pushing "k" key on a keyboard trigger the signal "k pressed". An application can use as many IISI as needed.

Because it is a virtual and invisible input, it allows to be easily set up at any place of the interactive space, making it possible to fit its characteristic such as position, shape or size, to specific spaces, users or any other attribute. It is unobtrusive and does not require any particular clothe or devices to be worn. It does not denature an ambiance because it does not occult anything in the interactive space. This can be desirable for artistic performances and interactive theaters for example.

Also, because it is an intangible spatial input there cannot be any physical contact so the potential danger of the input is null. In addition, it does not require any physical space so the interactive space can be used later for other purposes without requiring any physical effort to put the input away. It is also a good for hygienic purposes, and can be used in public space without creating any risk of infectious disease due to direct contact, and can be used in places that suffer from hygienic issues such as kitchens or factories.

2 Problem

Although an IISI has interesting features, it also has serious weak points. Due to its invisible nature, the user cannot directly see the area without the help of tertiary means. This property generates an important issue while trying to find the position of one specific input.

In previous work, we developed the collision detection technology, using a fast collision detection algorithm[1], then we created a system making use of several IISI. For each of them, a sound was associated making possible to play simple songs. However, during a presentation of this system, we have seen how difficult it was for people who desired to use it, to find the precise position of these sensors and properly interact with them. Generating a simple song resulted in a very challenging task.

In this work, we are interested in exploring ways to effectively design systems that use IISI, by exploring solutions that can help the user to locate accurately such input. We focus on understanding of how participants can effectively locate an invisible and intangible 3D point in space, and what characteristic must be used to design a proper application of IISI, even in a cognitively demanding context.

3 Previous work

Reaching intangible objects is a major challenge that is difficult to handle because of the lack of feedback inherent to the intangibility of the input. Without touch, the user can no looger feel surroundings. Notable work can be found in the field of virtual reality (VR) as the user should be able to interact with his virtual, and thus intangicle, environment. Poupyrev, et al.[2,3] decomposed the interaction process in 2 broad categories: egocentric and exocentric. In exocentric interactions users interact with virtual enviroments (VEs) from the outside (also known as the God's eye viewpoint). It can be materialized by iconic version of the VE[4] or automatic scaling[5]. In the egocentric interactions, the user can carry out indirect selection through techniques such as ray-casting[5,6], or by simulating the presence of human part in the virtual world such as using the Simple Viruual Hand technique[7]. Out of the field of VR, a Fresnel lens-based prototype of intangible display has been created in order to investigate whether visual and audio feedback helped improve user performance in target acquisition of an invisible and intangible input[8].

Study about strategies for the exploration and the collection of spatial information about a new area is an important topic in the fields of cognitive psychology and behavioral geography. Most of studies in these areas concern issues about spatial cognition-how human beings deal with issues concerning relations in space, navigation and wayfinding. These strategies are based on the use of different perceptual information. Researches on the use of VE apparatus also explore multimodal output—visual and haptic or visual and audio[9]. However, the research results show that the additional audio or haptic outputs did not increase the subjects' spatial knowledge. It might increase their spatial knowledge only after special training on how to use effectively the other senses. For example, research on the implementation of haptic technologies within VEs has reported on its potential for supporting the development of cognitive models of navigation and spatial knowledge[10-12]. Nerveless the dominant channel used by sighted people is the visual channel[13].

4 User experiment Ⅰ

We firstly conducted an experiment aiming to test the impact on locative performance when using helper that reduces the number of dimension to estimate. A point in Euclidian space has 3 dimensions.

When a participant is asked to locate a specific point in space, we made the hypotheses that he will more naturally try to use marks that lie originally on the surrounding environment (floor, ceil, walls around...) in order to establish spatial relationships instead of identifying the exact position in the space. This hypothesis emphasizes on the importance of the use of external marks and relative distances to process the locative tasks. We also wanted to see if relationship between the input and our body could be effectively used.

4.1 Task

Twenty-seven participants were recruited, 13 females and 14 males, ages ranging from 21 to 32. They were mostly undergraduate and graduate students from engineering major. They all reported a daily computer use of at least 1 h, and 16 of them more than 5 h. They all had no vision problem and were able to perform the test. At the end of the session they received a compensation of about US $10.

Participants were asked to find an input taking form of a cube of 10-cm edges, in a 7-m long by 6-m wide and 2-m high space. Each test started by showing to the participant his starting position. This position has been carefully chosen regarding its distance to the location of the IISI (close, average or far) and its position relatively to the interactive space in order to give varying point of view for each test. Once the participant reached its starting position, the evaluator indicated the IISI position to locate in the space. After that, the participant was asked to concentrate on its location as long as needed. When ready, he had to turn back for 2 s in order to lose direct sight of the target and give time to the evaluator to step back. Lastly, the participant was asked to do one of these 2 types of task.

1) Reaching task

The participant had to place one of his fists at the location he believed the IISI was, without any feedback to help him. This task focuses on precision.

2) Finding task

In this second task, the participant was asked to find the input with his fist as fast as possible. To do so, a feedback on a screen was signaling when the input was found, meaning a collision between the input and any part of the body of the participant occurred. However, to be a successful test, the hand, or at least a part of the hand, of the participant had to be inside the volume that characterized the input. If the participant was out, or if another part of his body activated the input, the task was considered unsuccessful. Fig. 1 shows the possible cases. This task focuses on speed.

Fig. 1 Result frames from one of the camera used during tests. From this point of view, the picture on the left represents a successful test, while the others are unsuccessful cases

4.2 Experimental design

Now we will describe the 5 different setups of helpers schematized in Fig. 2. The first one, noted setup N, does not use any kind of helper. In that case, the user must find the input only based on their observation. In the second one, noted setup V, one visual marker, a squared label of 10-cm sides, is placed on a wall. If an imaginary line is drawn from the marker, perpendicularly to the wall in which the marker is hooked, the input is positioned somewhere on this line. Thus it was asked to the user to use this marker to help finding the input with more precision. Such marker naturally constrains 2 dimensions, the height (Y axis) and one of the other (X, or Z, depending on the wall it is hooked on). The third setup noted setup V_m, uses 2 visual markers: 1 on the wall similarly to the pervious setup, and 1 situated on the ground. These 2 markers constrain all 3 dimensions, and the input is located as the intersection of the 2 imaginary lines formed by the 2 markers, perpendicularly to the surface they are hooked on. For these 2 previous cases, the distances, from the wall on which the marker was hooked, to the input, were varying from 1.4 m to almost 10 m. In the fourth setup, noted setup H, we adapted the input height to the height of the participant. We used the height of their shoulders or knees. As it naturally constrains 1 dimension, the input was somewhere in the plane formed by "Y = participant' shoulder (or knee) height". Lastly the fifth setup, setup H_m, is like the fourth setup with an additional marker on the ground that constrains the last 2 dimensions. Hence the input was located at the intersection of the plane formed like in the fourth case, and the line generated by the marker on the ground. For each setup, each type of task (finding and reaching tasks) was performed 2 times, leading to a total of 20 tests per participant.

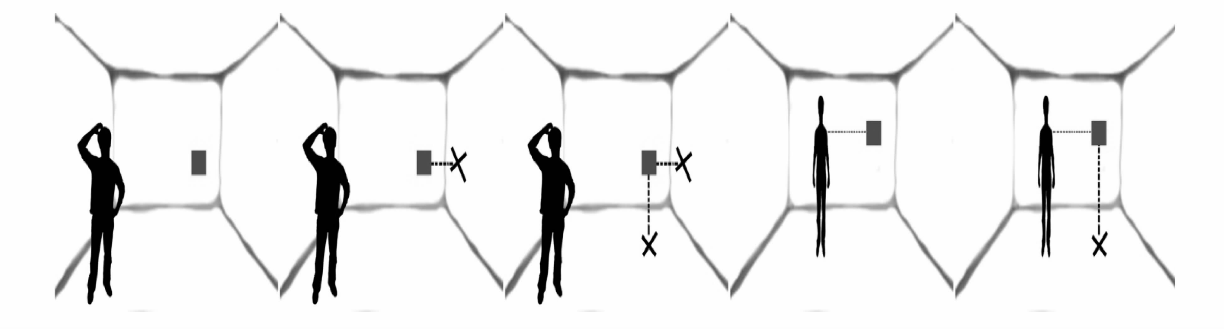

Fig. 2 Setup description. The red square symbolize the input location to be found. Setup N without helper. Setup V with a maker on the wall. Setup V_m with 2 markers, 1 on the wall and 1 on the ground. Setup H with the input height equal to the participant shoulder's height(for example). Setup Hm similar to previous with an additional marker on the ground

4.3 Results

Fig. 3 shows the average distance error in millimeters from the center of the input and the location pointed by the participant, in all the 5 setups during the reaching task (RT). The setup without any markers shows that when a location has been observed for some time, user can come back to this location with an average of 19.65-cm error distance. From that figure, we can see the results of 2 main groups: V/H and X/X_m, X being whether V or H. For V/H comparison, with improvement against N of 17.74% and 63.94% from respectively Setups H_s and H_m, we can see that the H group perform better than the V group which show −0.94% and 50.16% of improvement for respectively setups V_s and V_m. The case of V_s is very surprising for it shows no improvement at all for the RT. Secondly, we can see that X_m group shows an extremely significant improvement from "X".

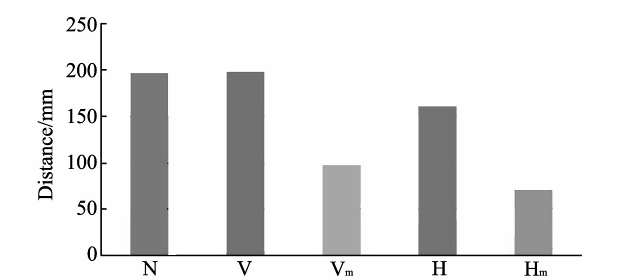

Fig. 3 Average distance error (in mm) from the 5 different setups for the reaching task

Fig. 4 focuses only on the Y axis (the vertical axis). On this figure we can see the significant improvement on this particular axis for the group H while the V group only shows really slight improvement. We can clearly see from the Setup H which only used the participant body's height, that this helper fills effectively its role which was to constrain the Y axis.

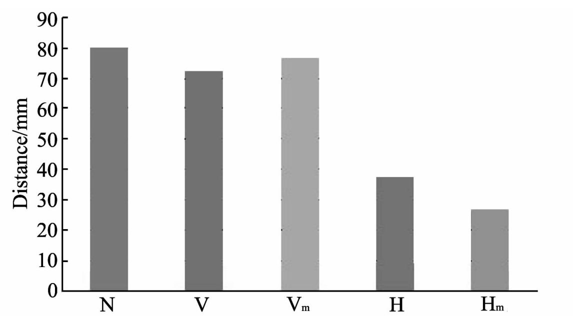

Fig. 4 Error on the Y axis for the 5 setups

Table 1 shows results gathered during the finding task(FT). The first row of results present the average time needed by participant in order to find the input. The second row of results present the success rate of the task. The task is defined successful if and only if the input was found with the hand in a time range inferior to 20 s. Not surprisingly, the longest time and lowest success rate are both under the setup N. Setup V shows actually better result than setup H, but it displays a low success rate. The addition of the marker on the ground, for each group V/H has a significant impact on the success rate (97.5% for setup V_m and 100% for setup H_m). In regard to the timings, we can see the relatively low result for the setup H weight against the other setups beside setup N. Also, setup H_m confirms being the most effective in all cases investigated here, with a 45.99% diminution of the average time needed to complete the task from setup N.

Table 1 Finding task results

	N	V	V_m	H	H_m
Average time/s	8.434	6.541	6.575	7.869	4.555
Success rate/%	53.40	58.50	97.50	71.80	100

5 User study Ⅱ

The first experiment was an important step toward understanding basic knowledge necessary to design interactive system based on IISI. However, the study took the form of a psychological experiment and cannot give any certitude in a real interactive context. The objective of this second user study was to stress this input technique through a challenging context of use. To do so, we chose to develop a dancing game inspired by dance dance revolution (DDR). To know more about the whole phenomenon concerning this game, we invite the reader to consult Refs. [14] and [15]. During a party of this game, the user must follow the steps represented by arrow on a screen, by stepping on the corresponding arrows on a mat under him. The steps are matched with the tempo of a song that is played at the same time. In the original version, there are 4 different steps: up(ahead of the user), right, left and down(behind the user). This game is extremely demanding cognitively and requires a lot of concentration from the user. Listening to the music while monitoring arrows on the screen and moving the body at the right location on the right tempo is very challenging. We believe the accumulation of all these tasks cannot be done correctly if the input is not intuitive and usable enough. Hence, it creates a good environment to stress the input.

5.1 Interface and setups

Overall, our prototype is based on 9 inputs as shown in Fig.5. Four of them form the foot layers, they are placed in a similar cross-shaped pattern that exists in DDR. The four others form the hand layers. The cross pattern is again used. However, the gap between the inputs as well as their height is specific for each setup. The last input is called the "special". This input was placed at 2 m on the right, and slightly behind the central position of the participant. This special input is aimed to force the user to move out from the initial position and not let him stay static at a single location during the whole song.

Fig.5 Game spatial input configuration. Reds inputs form the foot layer, blue ones form the hand layer and the yellow one is the special input

The setup NI only contained visual markers on the ground. Four were used in the cross-shaped pattern and represented IISI in both foot and hand layers. Another one is used for the special. The height of the center of the inputs forming the hand layer was set at 1.70 m and the height for the center of the special was set at 1 m. The IISI were cubes with edges fixed at 25 cm. IISI at foot layers were square (also 25 cm).

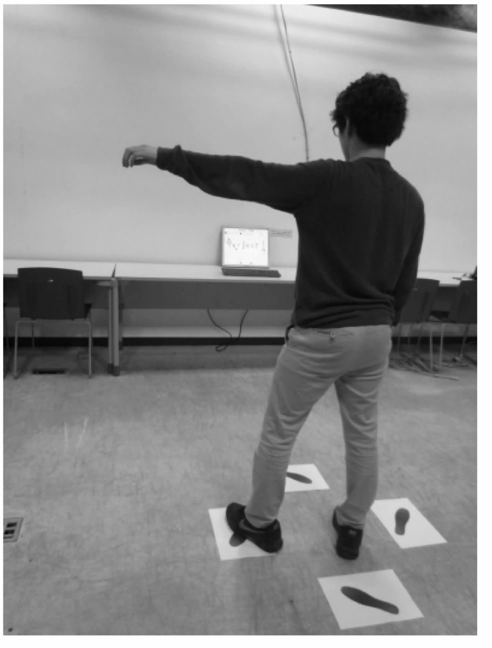

Fig.6 User testing the setup HI

The setup DI was based on the double markers principle of the setup V_m in the first experiment. It was the exact same one as the setup NI, however, we indicated the height of the hand layer's input by using 4 markers on the wall and the input size was not 25 cm but 20 cm to be consistent with results obtained when the experiment 1. Three markers were placed in front of the user so the same one indicated the up and down steps and the 2 others were used for left and right. These markers were placed at a distance of 3 m from the center of the layer. The special also had a marker placed on the adjacent wall at 2.5 m from the input.

The last setup HI, was based on the results obtained for the setup H_m. In this setup too, each input was above a visual marker on the ground. However, no markers on the wall were used. Unlike the 2 previous setups, this setup was re-adjusted for every participant, based on their size. Prior to the test, we measured the participants shoulder's height to adjust the height of the hand layers, hips' height for the height of the special, and arm length to adjust the gap between hand layers inputs. From the result obtained in the first experiment, we set the input size at 15 cm.

5.2 Task and experimental design

The 19 participants were 9 males and 10 females, ages ranging from 22 to 32. Most of them were undergraduate or graduate students in an engineering or media related field. They mostly had a low experience with DDR or any similar dancing game, as 6 of them have reported "know but never tried" this kind of game, 10 "just played few times", 1 "played time to time", 1 "plays regularly" and 1 "plays really often".

Participants played the game 2 times, one time using HI setup, and one time using alternatively NI or VI setup. As setup H_m gave a lot better results than other setups during the first experiment, we aimed to particularly stress the related setup within this cognitively demanding context. The song and steps that we used were the same for each try. A person with a long experience of playing DDR, tried the system rated the song as "difficult". The song was "do not stop the music" by Rihanna, and was known by almost all participants. We made a total of 50 steps, 25 at foot layer, 16 at hand layer, 2 specials, 3 doubles (2 input must be reached at the same time at the same layer), and 4 doubles with cross-layers (1 step at foot layer and 1 at hand layer at the same time). We positioned the arrows in order to follow the flow of the music, so they were not equally distributed resulting in some difficult parts and some easier ones.

For each step, we recorded if the input was activated in a certain range of time compared with the "true time". The closer the range was, the more points the participant earned. If we considered the "true time" for each step, we set up a timing window to 50 ms from the "true time" for "Perfect" rewarded by 1 000 points, another to 150 ms for "Good" rewarded by 500 points, the last one to 250 ms for "Bad" rewarded by 250 points. If there was no detection of input state at the end of the "Bad" timing window, the step was counted as "Miss" and no points were granted.

Prior to the participant's first try, the overall game rules were explained and the evaluator was in charge to demonstrate a try. After that, the concept of the first setup was explained. The participant could try for 1min to move and use the interface, while the GUI indicated each detected collision with any IISI. After making sure the participant understood everything, the first song using the first setup started. The same process was done again for the new setup used in the 2nd try.

5.3 Results

For the results, we used 2 types of data: the average total score of the tries, and we also extracted the average score performed only with hand and special, and regrouped them in the term of "Arms". We did this because this last category is really representative of IISI, while the foot layer is more an extension of a mat. Fig.7 shows the average score per game. As we can see that HI is showing the best results with 16 013.15 as average total score per game including 5 460.52 for Arms. It is followed by DI which shows 12 722.22 average points per game and 4 500 for Arms (respectively 20.55% and 17.59% decreased performance from HI). NI with an average of 12 450 points shows very close performance with DI, however, we can note low performance from Arms score with only 2 750 (51.32% inferior to HI).

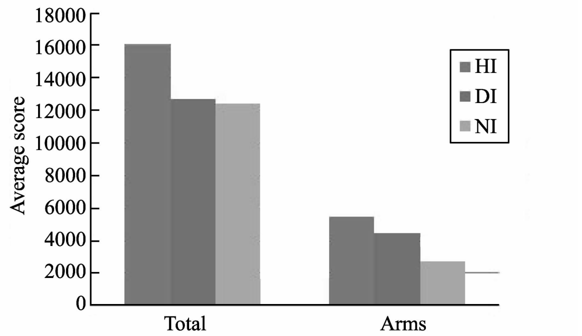

Fig.7 Average score per game for the 3 setups. On the left: the average Total final score; on the right: the average final score only counting hand layer and special

Fig.8 shows the results obtained by only considering the starting setup. As we changed at every participant, we wanted to see the difference in term of learnability and first contact with the interface. HI gets again the best results displaying an average of 14 000 points per game including 5 000 for Arms. The setup DI obtains a surprisingly low score of 9 250 points while N gets 11 100. However, on the Arms side, setup VI is the second with 3 300 and setup NI gets only 2 050.

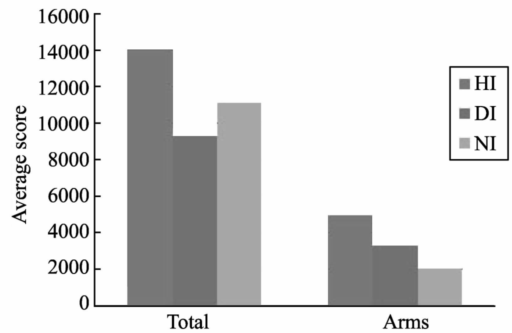

Fig.8 Average score obtained for each setup, only considering the first try

Fig.9 shows the difference between the setup used in the second try and the one in the first one. From the first to the second try, an improvement should naturally happens, because the participant acquired more experience of the system, knows better the song and its different movements, and can recognize and reproduce some patterns that he remembers from the first try. We can see that the switch from any interface from HI is very beneficial as it results in improvement by an average of 7 850 points, 3 175 for Arms. DI is also showing improved performance of about 4 312. 5 points, where 49% of them (2 125) arise from Arms. However, the setup N is showing very poor performances, with a loss of −1 200 points on average, and especially concerning Arms, with a substantial decrease of −2 400 points.

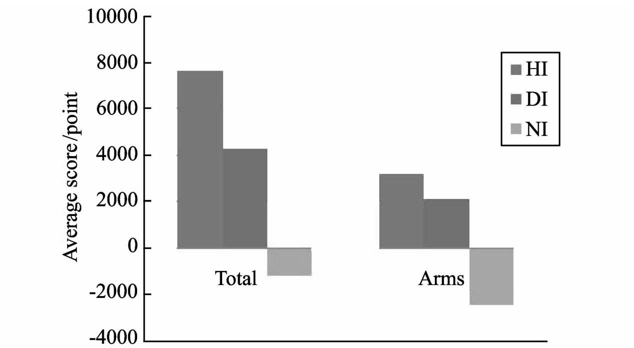

Fig. 9 Score difference between score obtained at the second try and the score obtained at the first try. The indicated setup corresponds to the one used for the second try

Fig. 10 shows the subjective rating in usefulness for each type of helper, given its setup. All setups had markers on the ground. However, only setup HI and setup DI had a second helper (respectively knowledge on the height and markers on the wall). Markers on the ground varies from 3.22(moderately used) for the setup DI to 4.1(often used) for the setup NI, while setup HI displays 3.84. However, the maximum and minimum grades come from the other helpers. The most useful helper, with a grade of 4.31, is the adjustment to specific height of the participant in the HI setup. However, the visual markers of the setup DI get only a grade of 2.33.

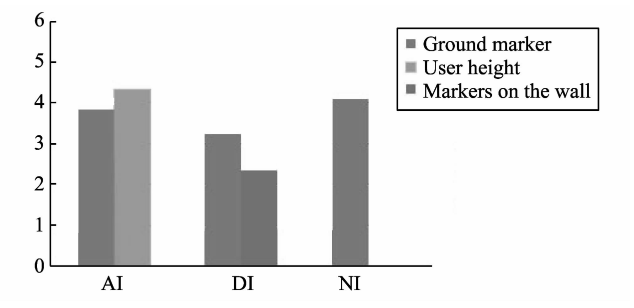

Fig. 10 Average usefulness rating of each type of helper for each setup. (1-Not used at all ; 6-Used all the time)

6 Conclusion

In this paper, we studied an alternative method of interaction in the context of whole body interaction, focusing on maximizing the freedom of interaction. We define freedom of interaction, analyze its impact and argue that interactive methods placing constraints directly on the user cannot be appropriate to maximize it. We propose to do a paradigm shift from body centered interaction to space centered interaction to displace these constraints to the space and investigate solutions based on invisible and intangible spatial inputs. As IISI are difficult to use because of their lack of sensible representation, we make 2 user studies that investigate the use of solutions that reduce the number of dimension to investigate. The results show that the locative task can be drastically improved by using markers on the ground, especially if the inputs heights have a direct relationship with the body of the user. Lastly, we have shown that even in cognitively demanding contexts, this design still works effectively. For future work we plan to deepen the study using IISI by testing their performance to control a display.

References

[1] Yoon T. Fast 3D collision detection algorithm between real-virtual objects using 2D intersection area and visual Hull. MS Thesis, HCI Lab, Soongsil university, 2009.

[2] Poupyrev I, Wegjorst S, Billinghurst M, et al. Egocentric object manipulation in virtual environment: empirical evaluation of interaction techniques. Proc. of Computer Graphics Forum, Eurographics Association and Blackwell Publishers Ltd, 1998,17(3): 41-52.

[3] Poupyrev I, Ichikawa T. Manipulation object in virtual worlds: categorization and empirical evaluation of interaction techniques. Visual Languages and Computing, 1999, 10(1): 19-35.

[4] Stoakley R, Conwaw M J, Paush R. Virtual reality on a WIM: interactive worlds in miniature. Proc. of the SIGCHI conference on Human factors in computing systems, New York, USA, 1995: 265-272.

[5] Mine M R, Brooks F P Jr., Sequin C H. Moving objects in space: exploiting proprioception in virtual environment interaction. Computer Graphics, 1997.

[6] Poupyrev I, Billinghurst M, Weghorst S, et al. The go-go interaction technique: non-linear mapping for direct manipulation in VR. Proc. of the 9th annual ACM Symposium on User Interface Software and Technology, Seattle, Washington, USA, 1996: 79-80.

[7] Sturman D J, Zeltzer D, Pieper S. Handson interaction with virtual environments. Proc. of the 2nd annual ACM SIGGRAPH Symposium on User Interface Software and Technology, New York, USA, 1989:19-24.

[8] CHAN Li-wei, KAO Hui-shan, CHEN M Y, et al. Touching the void: direct-touch interaction for intangible displays. Proc. of the 28th International Conference on

Human Factors in Computing Systems, Atlanta, Georgia, USA, 2010.
[9] Gunther R, Kazman R, MacGregor C. Using 3D sound as a navigational aid in virtual environments. Behaviour and Information Technology, 2004, 23 (6): 435-446.
[10] Witmer B G, Bailey J H, Knerr B W, et al. Virtual spaces and real world places: transfer of route knowledge, International Journal of Human-Computer Studies, 1996, 45: 413-428.
[11] Giess C, Evers H, Meinzer H P. Haptic volume rendering in different scenarios of surgical planning. Proc. of the 3rd PHANToM Users Group Workshop, MIT, Massachusetts, USA, 1998.
[12] Darken R P, Peterson B. Spatial orientation, wayfinding, and representation. Handbook of Virtual Environments Design, Implementation and Applications, Lawrence Erlbaum Associates, Inc., New Jersey, USA, 2002: 493-518.
[13] Heller M A, Schiff W. The psychology of touch. New Jersey: Lawrence Erlbaum Associates, 1991:23-60.
[14] Liu D. A case history of the success of dance dance revolution in the United States. [2011-08-11] http://www.stanford.edu/~dvliu/sts.pdf.
[15] Hoysniemi J. Games, user interface and performing arts: international survey on the dance dance revolution game. Computers in Entertainment (CIE) -Theoretical and Practical Computer Applications in Entertainment, 2006,4(2): 8.