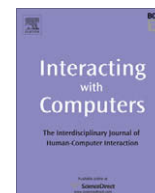


Contents lists available at [ScienceDirect](http://ScienceDirect)

# Interacting with Computers

journal homepage: [www.elsevier.com/locate/intcom](http://www.elsevier.com/locate/intcom)

## On scaling strategies for the full-body postural control of virtual mannequins

Ronan Boulic\*, Damien Maupu, Daniel Thalmann

VRLAB, Ecole Polytechnique Fédérale de Lausanne, Station 14, Faculté IC, CH 1015 Lausanne, Switzerland

### ARTICLE INFO

#### Article history:

Received 3 October 2008

Accepted 7 October 2008

Available online 14 October 2008

#### Keywords:

Full-body movement  
Full-body interaction  
Posture  
Reaching  
Scaling  
Virtual mannequins

### ABSTRACT

Due to its intrinsic complexity, full-body postural input has been mostly limited to off-line motion capture and to on-line puppetry of a virtual character with little interaction with its environment (e.g. floor). The motion capture technology is now mature enough to envision the on-line full-body postural control of virtual mannequins involved in precise reaching tasks. We have investigated such tasks for mannequins of differing body heights in comparison to that of the system user. Such broad-range avatar control is relevant for virtual prototyping in various industrial sectors as a single person is responsible for evaluating a virtual prototype for a full range of potential end-users. In the present paper we report on two scaling strategies that can be enforced in such a context of height-differing avatar control. Both scaling strategies have been evaluated in a wide-range reach study both in front of a stationary immersive display and with an HMD. A comparison is also made with a baseline scenario, which exploits a simple rigid shape (i.e. a proxy), to assess the specific influence of controlling a complex articulated avatar.

© 2008 Elsevier B.V. All rights reserved.

### 1. Introduction

Body movements have been shown to be a necessary condition for assessing the feasibility and the relative distance of a reaching task (Mantel et al., 2005). However, due to its intrinsic complexity, exploiting full-body postural input has been mostly limited to off-line motion capture (Menache, 1999) and to the on-line puppetry of a virtual character (Sturman, 1998) with little interactions with its environment (e.g. floor in Shin et al. (2001)). In the field of Virtual Reality, a few full-body avatar control attempts have been relying on the magnetic tracking technology (Badler et al., 1993; Molet et al., 1999). The execution of immersive 3D tasks such as reaching, grasping and manipulation have otherwise mostly been relying on isolated virtual tools driven by the locations of the user's hands (Mine et al., 1997). This approach has proven to be sufficient for a large range of applications in Virtual Environments (Bowman and Hodges, 1997; Ware, 1990; Ware and Jessome, 1988).

*Abbreviations:* HMD, head mounted display; FOV, field of view; IK, inverse kinematics; Strategy A, *reference* strategy with immersive display at scale 1/1 and egocentric reaching; Strategy B, *visuocentric* strategy with immersive display at scale 1/1 and scaled sensor data; Strategy C, *egocentric* strategy with scaled immersive display and egocentric reaching; the controlled entity factor has two levels, *baseline/avatar-control*; the reach difficulty factor has two levels, *free-space/obstructed*.

\* Corresponding author. Tel.: +41 21 693 52 46; fax: +41 21 693 53 28.

E-mail addresses: [ronan.boulic@epfl.ch](mailto:ronan.boulic@epfl.ch) (R. Boulic), [damien.maupu@epfl.ch](mailto:damien.maupu@epfl.ch) (D. Maupu), [daniel.thalmann@epfl.ch](mailto:daniel.thalmann@epfl.ch) (D. Thalmann).

In recent years, the motion capture technology has matured with the use of active optical markers (e.g. Phasespace) and some recent systems are much more affordable. The intrinsically enactive knowledge of full-body movements has long been serving for the sole virtual tool positioning. We believe that it can also be exploited at the full-body scale for the control of virtual mannequins interacting with their environment. It makes sense for those more demanding applications for which the full-body postural information is necessary to make decisions. It is now pertinent to envision the on-line full-body postural control of virtual mannequins involved in precise interaction tasks (ISPR Kraal et al., 2000). The goal of the present paper is to assess this interaction channel when reaching visible targets displayed on a large immersive screen. One key aspect of the study is to evaluate subjects' performance when controlling a virtual mannequin of differing body height. Such an ability is necessary in numerous industrial sectors to evaluate the use of a virtual prototype by a range of virtual mannequins representative of the target end-user population. In such applications, it is particularly important for an evaluator to experience the same difficulty to reach a given target as the virtual mannequin that he/she is currently controlling. As an extreme illustration we can imagine the case of a classroom designer evaluator who is radically taller than the target user population of kindergarten children.

With that ecological consistency requirement in mind, we have examined the effect of two scaling strategies (sensor scaling vs environment scaling) on the reaching response duration when controlling a virtual mannequin of differing body height. The reach

tasks were located at a constant distance in front of the participants, at four constant heights. A two levels “controlled entity” factor was used to assess the specific influence of controlling a complex articulated avatar (avatar-control level) compared to a simple rigid virtual shape (i.e. a proxy) representative of the state of the art (baseline level). An additional “reach difficulty” factor with two levels (free-space/obstructed) has been considered to highlight the sensitivity of the reaching response duration to the complexity of the environment. The experiment was designed in order to always ensure a feasible reach that can be translated into a measurable reaching response duration within a time-out limit. This consideration prevented the exploration of extreme body height differences such as the adult-child illustration mentioned above.

This paper is organized as follows: the next section recalls the prior efforts on full-body interaction and on characterizing reaching performances. Section 3 describes the rationale of using scaling strategies and discusses the viewpoint type issue (first-person/third-person). Section 4 recalls the key hypothesis being tested in the study while Section 5 describes the experimental protocol. Results are presented in Section 6 and discussed in Section 7 before concluding the study.

## 2. Related work

### 2.1. Full-body interactions

A large body of work has been dedicated to identify the most intuitive types of interaction with complex 3D Virtual Environments, especially for tasks such as reaching and manipulating virtual objects (Bowman and Hodges, 1997; Mine et al., 1997; Stoakley et al., 1995; Ware, 1990; Ware and Jessome, 1988). The matter is complicated by the fact that users immersed in a Virtual Environment have the tendency to perceive distances differently than in standard settings (Loomis and Knapp, 2003). Studies have also shown that the

estimation of reachability of a target depends on the posture prior to performing the reach. For example a secure seated posture leading to overestimation vs an unstable standing on one foot posture leading to underestimation (Gabbard et al., 2007). Mantel et al. have shown that ultimately, i.e., when no hint can help, it is necessary to move the viewpoint to determine the relative distance to a target of unknown size (Mantel et al., 2005). This strongly suggests that full-body immersion and movement are necessary for high quality 3D interactions. This in turn brings in the related issues of the best type of immersive display (Pausch et al., 1993), of viewpoint (Schaffer and Bowman, 2004), and whether displaying the virtual body can help or not (Draper, 1995).

Recent technological progress in active optical motion capture (e.g. Phasespace) now allow real-time full-body interactions with virtual environments (Peinado et al., to appear). Such an ability is necessary for those demanding applications requesting the control of virtual mannequins for the evaluation of virtual prototypes (Kraal et al., 2000). One critical aspect to ensure when an evaluator controls a virtual mannequin of different body height is to replicate that mannequin’s egocentric reach perception as it clearly depends on its body height (Caird, 1994; Carello et al., 1989).

### 2.2. Reach performance

In (Boulic et al., 2006) we have described an experimental study to quantitatively assess the relationship between the reaching duration and ten levels of normalized target heights. Subjects were standing at a distance of 0.3 body height from a large screen (Fig. 1a); the goal was to successively reach a target displayed at discrete heights that were expressed in body height units, from 0.2 to 1.1 body height (Fig. 1b and c). Despite large inter-subject differences among absolute reach durations (Fig. 2a), much smaller intra-subject variations per target height was observed (Fig. 2b). A minimum of normalized reaching durations is observed around 0.7–0.8 body height for that experimental context (Fig. 2c).



**Fig. 1.** (a) The subject is standing in the default start posture before reaching a target located at a normalized reach height (= absolute reach height/body height) on the screen, (b) reaching posture for a high target, (c) reaching posture for a low target.

### 3. Context of the study

#### 3.1. Rationale for the choice of avatar control strategies

The study aims at quantifying the reaching response duration as a function of target heights for subjects fully immersed in a Virtual Environment while controlling the posture of an avatar. For that purpose the subject movement is exploited in real-time with the help of a set of active optical sensors (see Section 5.2.1 for more details). Three avatar control strategies have been considered:

- *Reference strategy (A)*: control of a same-height avatar. Both the visual display and the postural sensor input are respectively presented and exploited at scale 1/1. Its purposes is to calibrate the reaching duration as a function of target heights.
- *Visuocentric scaling strategy (B)*: control of a height-differing avatar by scaling the sensor input to match the body height of the avatar. In this context the subject has to rely on the immersive visual feedback to guide the avatar towards a reach goal expressed in absolute coordinates.

- *Egocentric scaling strategy (C)*: control of a differing-height avatar by inversely scaling the displayed environment, including the target and the avatar, so that the displayed avatar height matches the user height. This ensures the correspondence of egocentric spatial ability between the subject and the controlled avatar.

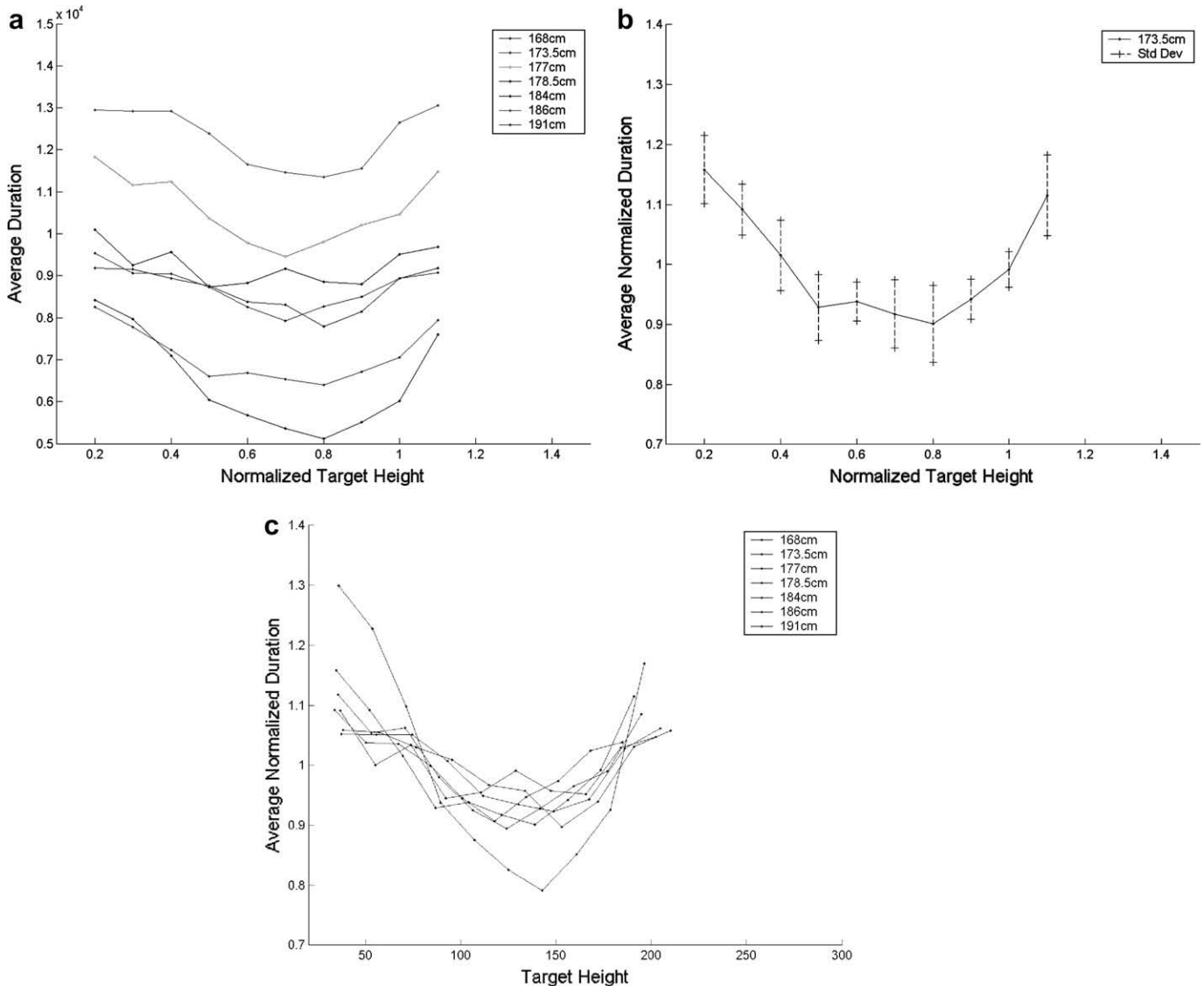
We now briefly compare the two scaling strategies B and C in terms of prior usage and condition of equivalence.

#### 3.1.1. Visuocentric strategy (B): scaling the sensor data

Given the full-body sensor data measured from the current subject posture, two dual approaches can be exploited for the on-line control of a height-differing avatar. The traditional approach in Animation (*Autodesk MotionBuilder*) and Virtual Prototyping for the industry (*Kraal et al., 2000*) is to scale the 3D sensor position data by the ratio  $R$  of the body heights, with:

$$R = \text{avatar\_height} / \text{subject\_height} \tag{1}$$

The scaled sensor positions can then drive constraints associated to the avatar skeleton with the posture reconstruction algorithm outlined in Section 5.2.1.



**Fig. 2.** (a) Dispersion of inter-subject reaching response durations, (b) average normalized reaching durations per target height with standard deviation for one subject, (c) all subjects' average normalized reaching duration; a minimum occurs around 0.7–0.8 body height.

Although the resulting avatar posture is basically the same as the subject posture when expressed in term of joint angles, this approach results in a scaled up or scaled down displacement of the subjects body parts (e.g. hands) in the Cartesian task space. Hence the reachable space is also scaled by the ratio  $R$  that brings the avatar hand to a different absolute location compared to the subject hand. Such discrepancy between the subject egocentric spatial ability and the one of the avatar becomes apparent in some viewing conditions (detailed further in the section). In such a case the participant has to rely on the visual feedback modality to adjust his/her posture when having to achieve a reach task expressed in absolute coordinates (similarly to a puppeteer). For this reason we qualify this strategy as the visuocentric strategy (Scheidt et al., 2005).

### 3.1.2. Egocentric strategy: scaling the virtual environment

The alternate scaling strategy to control a height-differing avatar is to scale the displayed environment, including the avatar and the targets, with the factor  $1/R$ . For example, in case the subject was an adult and the avatar a child, the virtual environment would appear enlarged as the ratio of subject height to avatar height is greater than 1. Most importantly such an approach preserves the subject egocentric perception of space.

### 3.1.3. Visual equivalence and difference of the scaling strategies

The two scaling strategies are mathematically equivalent when the visual feedback is computed for a viewpoint collocated with the avatar eyes and displayed to the subject with an HMD. We name this viewpoint the avatar first-person viewpoint. In such a context the target has the same image coordinates for both strategies as shown by the point target in Fig. 3a. In this figure the subject is taller than the controlled avatar; as a consequence the sensor data are scaled down in the visuocentric strategy B (Fig. 3a-left) whereas the virtual environment is scaled up in the egocentric strategy C (Fig. 3a-right), both being visually equivalent.

On the other hand, if another type of viewpoint and/or projection is exploited, the target image coordinates may differ for the scaling strategies as illustrated for an orthographic projection in Fig. 3b (visuocentric on the left and egocentric on the right). The next section justifies the viewpoint and projection choice.

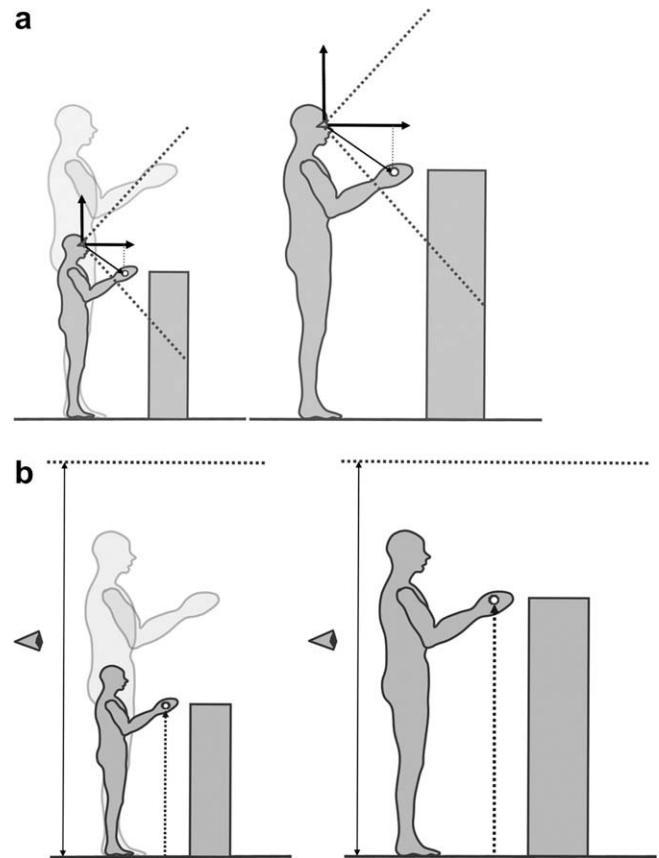
## 3.2. First-person viewpoint vs third-person viewpoint

The standard approach in industry is generally to exploit the visuocentric strategy because it is simpler to scale the sensor data than to scale a complex virtual prototype (Kraal et al., 2000). The associated avatar first-person visual feedback is usually displayed with a perspective projection in a HMD.

However, in our case, the HMD limited FOV (Fig. 4) allows the subjects to see only a fraction of the planned targets that are distributed at four heights from 0.5 m to 1.85 m (Fig. 8). As a consequence a search phase is sometimes necessary for the lowest target. This, together with additional comfort issues, bias the measured reaching durations towards longer values when reaching low targets as already reported by Draper (1995). For these reasons, the study is structured in two successive sessions, first with a third-person viewpoint (more detail in Section 5.2.4), then with a first-person viewpoint (Fig. 4).

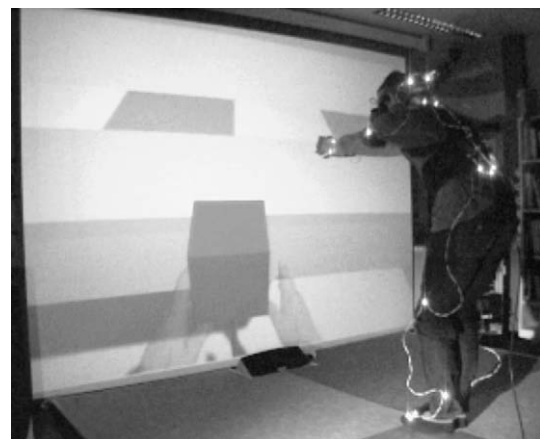
### 3.3. Common aspects of the three strategies with a third-person viewpoint

The choice of a third-person viewpoint induces the non-visual equivalence of the scaling strategies (Figs. 3b and 5) summarizes the remaining invariant characteristics of each pair of strategies for the avatar control experiment:

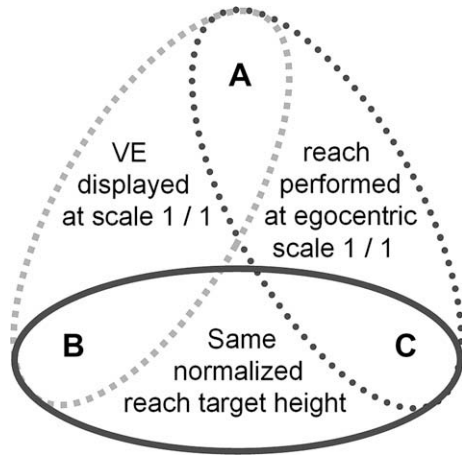


**Fig. 3.** (a) The point target appears at the same image coordinates for both scaling strategies with an avatar first-person perspective projection, (b) the target image coordinates do not coincide with an orthographic projection.

- *Same visual display of the Virtual Environment at scale 1/1 (ellipse AB):* only the sensor data are scaled in the visuocentric strategy (noted strategy B) to control an avatar with a different body height. The strategy B shares the same visual feedback of the virtual environment with the reference strategy A.
- *Same egocentric spatial reaching ability at scale 1/1 (ellipse AC):* only the visual feedback is scaled in the egocentric strategy (noted strategy C). Subjects can therefore experience their egocentric reaching ability in both strategies A and C.



**Fig. 4.** The subject sees with the HMD only a very limited portion of the reachable space intended to be tested in the full-body reaching experiment ; in this snapshot the first-person view is also displayed on the large screen for assessment purposes by an external operator.



**Fig. 5.** Common characteristics of the three tested strategies; (A) same-height reference, (B) visuocentric strategy (scaled sensor data), (C) egocentric strategy (scaled environment).

**Table 1**  
Distribution of subject body heights (m).

1.63	1.63	1.67	1.68	1.71	1.80	1.81	1.82	1.83	1.90	1.91
------	------	------	------	------	------	------	------	------	------	------

- *Same normalized target height (ellipse BC):* the scaling achieved in both strategies B and C ensures that the normalized target height (= target\_height/avatar\_height) is constant, hence subjects should produce the same final posture.

**4. Tested hypothesis**

The experiment is structured to assess the influence of four factors on the reaching performance when executed in Virtual Environments:

**Viewpoint type:** We aim to confirm the prior findings that first-person and third-person viewpoint types are not equivalent for low target reaches. The corresponding null hypothesis is the production of the same reach performance for both viewpoint types.

In line with the previously mentioned visibility issue with the first-person viewpoint (Section 3.2), the following hypotheses assume only measurements obtained with the third-person viewpoint modality:

**Controlled entity and reaching difficulty:** we want to assess the influence of the type of controlled entity (baseline/full-body avatar) for the control of a same-height entity (reference strategy). The reaching difficulty factor allows to study both the simplest possible reaching behaviour (free-space) and a more realistic one for real-world applications (obstructed). By construction, the obstructed difficulty level should lead to longer reaching durations. The null hypothesis is thus specified independently for each difficulty level, i.e. that the same reach performance is produced for both types of controlled entities. We also analyze the normalized reaching characteristics as this conveys useful information for the intended application field.

**Scaling strategy:** we want to assess whether the two scaling strategies allow to produce the same reaching behaviour as a person of differing-height. If the height distribution of height-differing entities is the same as the one of the participants, the corresponding null hypothesis would be that the reaching performance is the same for the reference strategy and each scaling strategy. If the height distributions differ, it is preferable to analyze the normalized reaching characteristics. The second hypothesis relates the two scaling strategies independently from the reference strategy. As they have been tested for the same-height distribution, it is valid to formulate the null hypothesis that they produce the same reaching performance.

**5. Materials and methods**

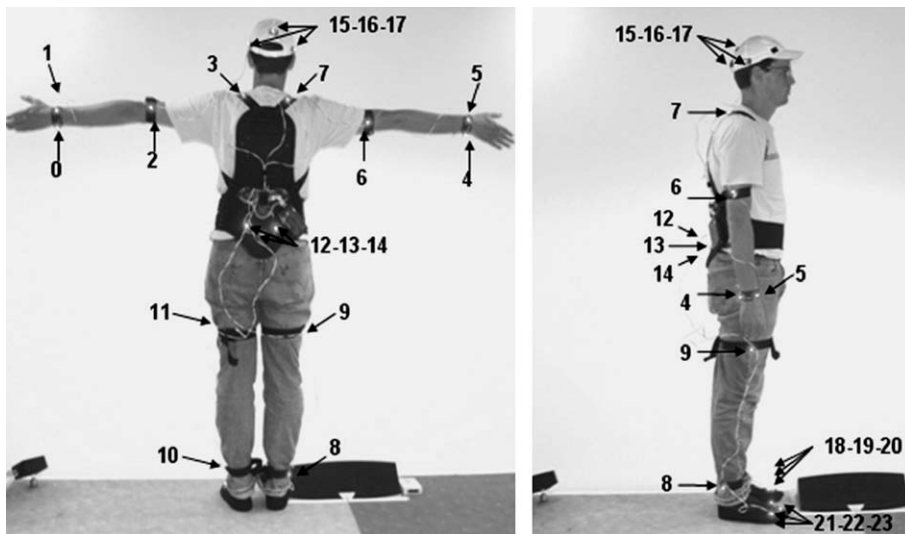
*5.1. Participants*

Eleven naïve subjects (two females and nine males, aged from 25 to 30) participated to the study. All had normal or correct-to-normal vision. None had counter-indication for standing-up over the duration of the study. They all gave prior written consent. The whole experiment lasted approximately 110 min, including regular compulsory seated pauses to prevent accumulated fatigue. The subjects body heights were distributed between 1.63 and 1.91 m, with an average of 1.76 m (Table 1).

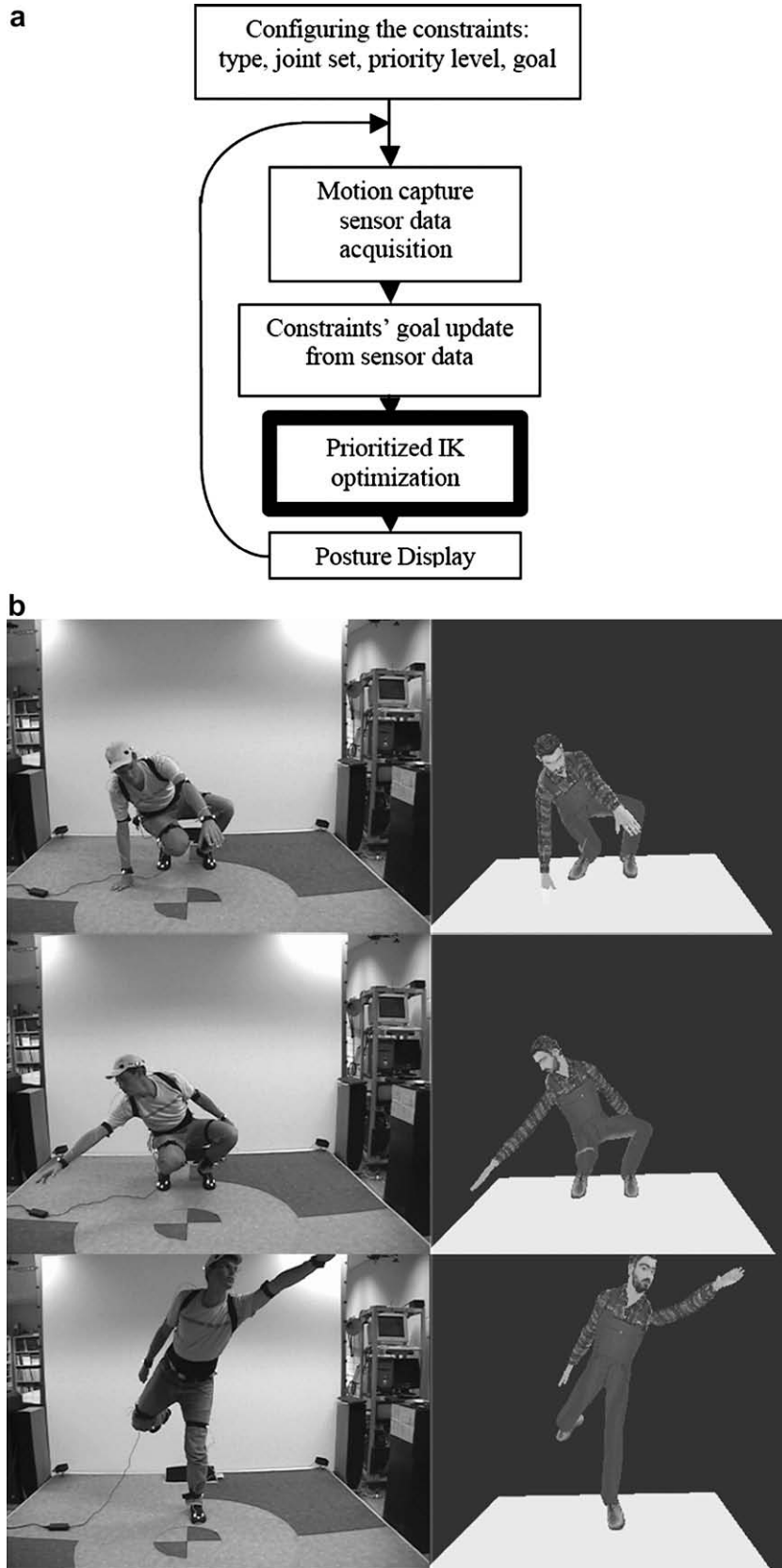
*5.2. Apparatus*

*5.2.1. Full-body motion capture*

The capture of the participant posture and the reaching tasks has been designed to be equivalent for right- and left-handed participants. We rely on the Phasespace active optical motion



**Fig. 6.** Configuration of active markers used for the full-body avatar control for simple reaching tasks.



**Fig. 7.** (a) Architecture of the on-line posture reconstruction, (b) real and reconstructed postures.

capture system to obtain the real-time 3D location of a set of 24 LED markers placed on the subjects as follows: abdomen and torso (5), head (3), each arm (3), each leg (2), each foot (3)

(see Fig. 6). This configuration allows to capture the full-body posture except the wrist flexion as we focus on a simple reach task without manipulation. On the other hand it provides the

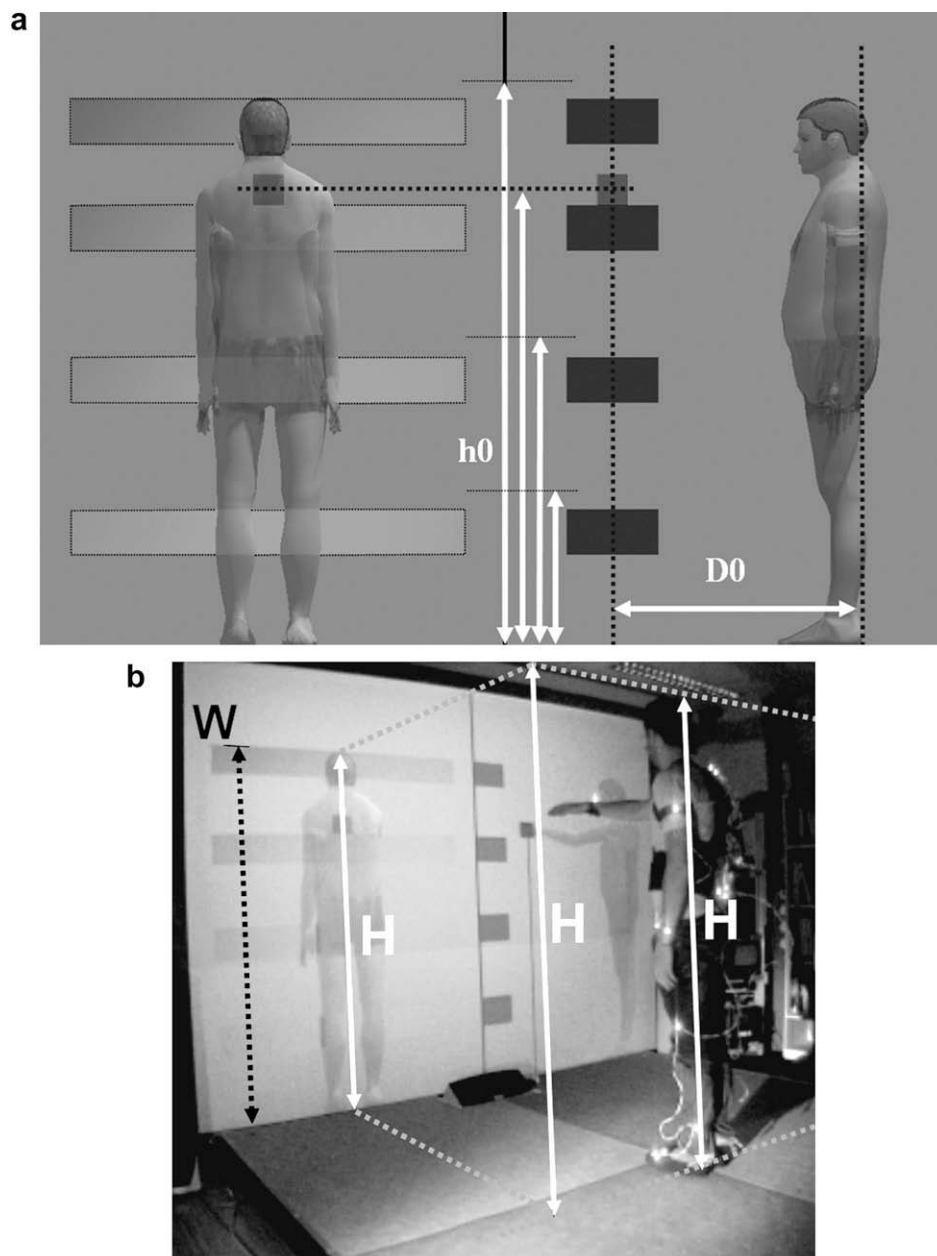
position and orientation of the foot tip in case of standing on tiptoe for difficult reaches.

The active optical technology is able to identify the markers in real-time, even after being temporarily occluded during the interaction. In case of occlusion the last known position of the marker is re-used by the reconstruction module. This feature allows to continuously reconstructing the full-body posture from the knowledge of the marker locations. The reconstruction approach we have adopted is based on a numeric inverse kinematics solver allowing to associate priority levels to postural constraints. This allows to enforce important properties first (e.g. feet stay on the ground) while less important adjustments are made in the remaining solution space (Baerlocher and Boulic, 2004). The overall reconstruction algorithm works as illustrated in Fig. 7. On average, each iteration of the posture reconstruction costs about 6–9 ms. However, despite this good performance, the numeric

solver limits each postural variation to a maximum norm to guarantee the stability of the linearized solution. If the position error is large the solver incrementally converges towards the optimal solution. This results in a form of low-pass filtering of the subject movement that may slightly slow down the interaction. For this reason the subjects have been instructed to perform the reaching tasks at a regular and normal pace. Detailed information on the posture reconstruction architecture can be found in (Maupu et al., 2007).

### 5.2.2. Avatars

Available avatar body heights were ranging from 1.56 to 1.93 m with about 0.1 m intervals. They were modelled according to the methodology from Kasap and Magnenat-Thalmann (2007). Each subject controlled one same-height avatar and a one height-differing avatar that were selected as follows:



**Fig. 8.** (a) Combination of fixed viewpoint and orthographic projection showing the target in front of a set of shelves used as a reference decor for the scaling, (b) immersive display for the reference strategy (same-height avatar).

- For strategy A, the avatar closest in height to the user was chosen. A small scaling was applied to the 3D model to have a perfect body height correspondence with the subject height.
- For strategy B, the height of the height-differing avatar was determined by the experimenter so as to keep the reaching tasks always feasible for all subjects. For this reason we used the following subset of height-differing avatars: 1.65 m, 1.76 m, 1.82 m. Six participants controlled a smaller avatar and five controlled a taller one. The average height of this height-differing set was 1.72 m.
- For strategy C, only the virtual environment was scaled; the user and avatar heights remained the same. Thus C used the same avatar as in strategy A.

### 5.2.3. Virtual environment used for the reaching tasks

A virtual environment made out of horizontal shelf-like volumes was used and displayed differently depending on the reaching difficulty factor. This factor can take two levels, either “free-space” when displayed in the background (Fig. 8a), or “obstructed” when placed at the same distance as the target. This virtual environment served two purposes: (1) to highlight the sensitivity of the reaching response duration to the complexity of the environment, (2) to provide a permanent visual reference of the four target heights across the scaling strategies.

For all reach tasks performed at scale 1/1 (strategies A and B), the avatar stood at a distance  $DO = 0.5$  m from the target projection on the floor.  $DO$  is measured from mannequin’s heel to the target projection on the floor. Targets were displayed at the following absolute heights  $h0 = 0.5, 0.95, 1.4,$  and  $1.85$  m (Fig. 8a).

Two values of  $h0$  were chosen in easy-to-achieve mid reach heights and two others were selected at difficult-to-achieve lower and higher heights. However, for comparisons to be possible among subjects and scaling strategies, the range of  $h0$  values had to be feasible by all subjects even after the 1/R environment scaling applied to both  $DO$  and  $h0$  for strategy C. The same reaching feasibility requirement led to limit the height difference between the subject and the controlled avatar as mentioned in the previous section.

### 5.2.4. Immersive display and projection

We exploited the MVISIO multi-device graphic engine for the rendering and the skinning of the avatar (Peternier et al., 2008). We adopted a constant combination of orthographic projections as shown in Fig. 8. The projection parameters were chosen to ensure an immersive display at scale 1/1 for the subject. We selected a combination of two viewpoints and orthographic projection allowing the participant to determine without ambiguity the target 3D location. The combined images were displayed on a large screen ( $3\text{ m} \times 2.2\text{ m}$ ) as shown in Fig. 8b for a reach case of a 1.4 m high target with a same-height avatar. The avatar was also partly transparent to allow see-through, hence fulfilling our requirement of always visible target and equivalent apparatus for right and left-handed subjects. To conclude this section on the provided visual feedback the rigid virtual shape used in the baseline context is illustrated in Fig. 9.

## 5.3. Procedure

### 5.3.1. General organization

The combination of the two factors, respectively “controlled entity” with baseline/avatar-control and “reach difficulty” with free-space/obstructed, produces four conditions (Fig. 10 top layer); each characterizing the spatial configuration in terms of appearance and relative position of the target, the virtual environment and the controlled entity. For each condition the three scaling strategies have

been tested, respectively the same-height reference strategy A, the visuocentric strategy B and the egocentric strategy C (Fig. 10 bottom layer). As a consequence the participants experienced twelve different types of trials according the combinations of spatial configurations and scaling strategies.

### 5.3.2. Experimental design

Each session started with a 30 s initial training phase for each of the four spatial conditions; the participant had to reach one target with the reference strategy. Then, the four spatial conditions were performed in random order. A brief written qualitative feedback was gathered after each spatial condition mostly to ensure regular seated rests.

In each spatial condition, the three scaling strategies were performed in random order, with an initial 30 s training phase per strategy. The subjects were not told about the nature of the currently enforced strategy.

Each strategy consisted of six reaching tasks at the four heights and in random order too; the lowest and highest reach heights were performed twice. The reaching behaviour was evaluated through the measurement of each reaching task duration. A reaching task was always initiated from the standard rest posture illustrated in Fig. 8a. A sequence of three audible signals (i.e. a “beep”) notified the participant that a new reach would soon start. The timer then started when the new target appeared on screen and automatically stopped when the system detected the intersection of the target box with the hand proxy shown in Fig. 9. When the reach was successful, the target turned from red to green to provide a visual feedback to the subjects. The same detection process was used for both the baseline and the avatar control; in this latter case the hand proxy was collocated with the avatar hand but not displayed. There was an 8 s pause from the moment a reach was completed to the time a new reach was proposed.

Each of the four spatial conditions was tested 18 times for a total of 72 reaching tasks. The session duration was approximately 50 min for the third-person viewpoint. It was followed by a 5 min seated rest. We then conducted a similarly structured session in the first-person viewpoint but with a different random order of the spatial conditions, strategies and target heights.

## 6. Results

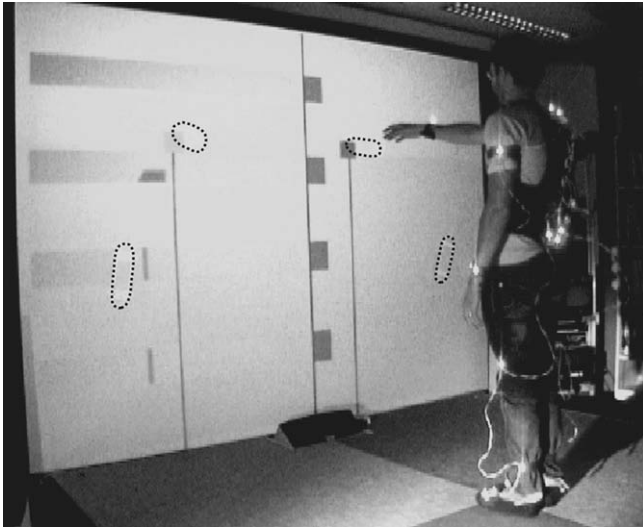
### 6.1. Outlier detection

Each viewpoint session produced 792 measurements. For the third-person viewpoint, the reach duration responses were distributed according to a normal law with a slight asymmetry with a longer tail above the general mean. It was reduced by removing a small number of time-out outliers (they were due to the lack of maturity of the numeric IK solver used to reconstruct the posture in real-time). In concrete terms we have retained the 767 measurements below a time-out value of 9 s, leading us to reject 25 measurements (3.16%) that were more than two standard deviations above the general mean of 3.69 s (stdev = 2.34 s).

### 6.2. Viewpoint factor

By construction of the experimental protocol, the first-person viewpoint session was evaluated after the third-person viewpoint session for all subjects. If a training effect were to be expected, it should have led to shorter reaching durations for all target heights. Instead, we have observed longer average reaching durations for the lowest 0.5 m target with the HMD: 62% longer in the (baseline,





**Fig. 9.** The baseline context consisted in achieving the same reaching tasks with a rigid virtual object collocated with the hands (the rigid shapes are outlined with the dotted lines).

using the first-person viewpoint with the HMD leads to longer reaching response durations when reaching low targets compared to the third-person viewpoint used in the present study. The remainder of the results section exploits only the measurements obtained with the third-person viewpoint session.

### 6.3. Controlled entity and reaching difficulty factors

We first analyze the spatial conditions when tested with the reference strategy. This strategy is characterized by an immersive display at scale 1/1 and an egocentric reaching ability (Fig. 11a).

#### 6.3.1. Comparison of absolute reach durations

The average of absolute reaching durations of all subjects is computed per target height, controlled entity and reaching difficulty. Mean values with one standard deviation are reported in Fig. 12a for the free-space reach and on Fig. 12b for the obstructed reach. The average durations increase for the extreme reaching heights compared to the two middle heights; this is consistent with the longer path to reach the targets and with the prior results shown in Fig. 2.

There is no significant difference among the average free-space reaching durations between the baseline and the avatar control levels for the 0.95, 1.4, and 1.85 m target heights. However the avatar control displays a 35% longer average duration for the 0.5 m target height ( $F(1,41) = 5.96, p = .019$ ).

For the obstructed level, the highest average durations are obtained for the 1.85 m target instead of the 0.5 m target for the free-space level. This can be explained by the shelf being more on the hand path of the highest target compared to the lowest target (Fig. 10). Longer average durations between the avatar control and the baseline are significant for each individual target height: respectively 39% for 0.5 m ( $F(1,41) = 6.40, p = .015$ ), 39% for 0.95 m ( $F(1,20) = 5.09, p = .035$ ), 57% for 1.4 m ( $F(1,20) = 8.28, p = .009$ ), and 32% for 1.85 m ( $F(1,39) = 4.53, p = .039$ ). Variances clearly increase as well (Fig. 12b).

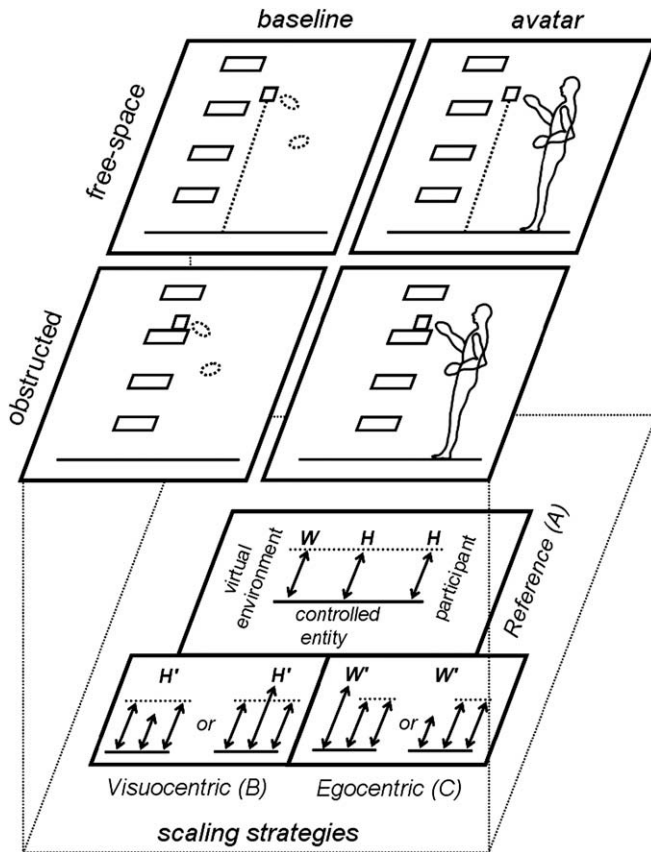
#### 6.3.2. Reaching duration normalization

As already noted by other authors and observed above, reaching durations were very different among subjects. We therefore normalized the absolute reaching response duration with the average duration per subject and per individual reaching series, i.e. for one spatial condition (controlled entity, reaching difficulty) and one specific strategy (= average of 6 reaching durations). The normalized duration of a reaching task is the ratio: *absolute reach duration/average duration*.

#### 6.3.3. Comparison of normalized reaching characteristics

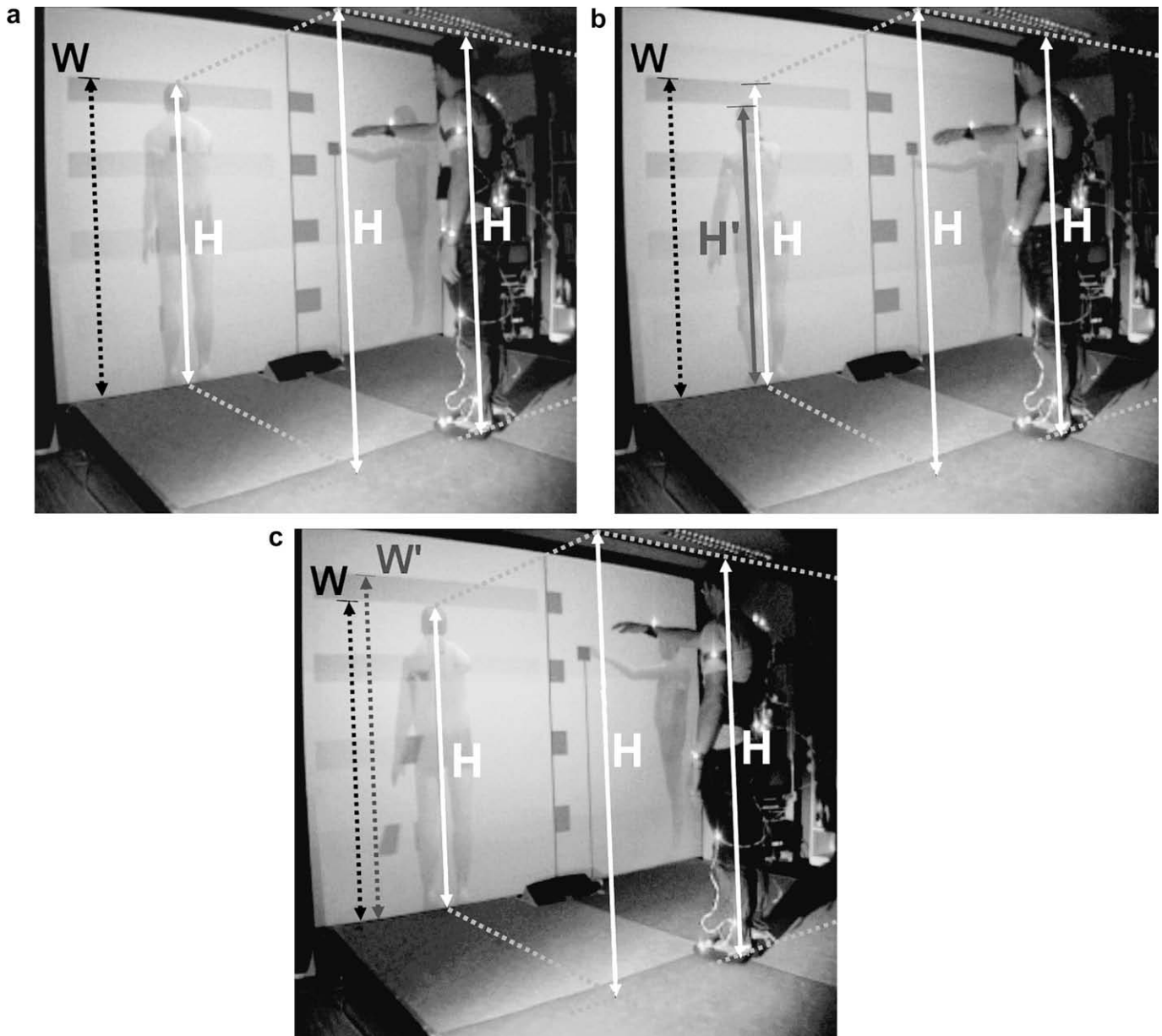
We have also established the reaching characteristic displaying the normalized duration as a function of the normalized height of the target (= *absolute target height/controlled avatar body height*). We have retained to fit a second order polynomial as being the simplest model close to the observed U shape in Fig. 2 and to the duration distribution observed in Fig. 12. Fig. 13 presents the two characteristics of free-space reaching for controlling a same-height entity (baseline or avatar). Please note that using the normalized height induces the data spread along the x-axis as the four target heights are divided by the eleven subject heights.

The quality of fit ( $R^2$  values) of the baseline control (0.35) is lower than the one of the avatar control (= .53) due to a greater dispersion around the model (i.e. larger residual Sum of Squares). This possibly comes from the ease of the baseline control. The added cognitive difficulty of the articulated avatar control led to longer absolute durations with less dispersion around the fitted curve (Fig. 13b). It is interesting to note the greater difference between the model minima and both sides extrema; e.g. the lowest target



**Fig. 10.** Illustration of the four spatial conditions (top layer), each evaluated with the three scaling strategies (bottom layer). The reference strategy A is characterized by a controlled entity having the same body height H as the participant and by a displayed environment W at scale 1/1. In the visuocentric strategy B the controlled entity H' is either smaller or taller compared to the participant whereas in the strategy C the environment W' is scaled by the factor 1/R.

free-space) condition ( $F(1,41) = 19.50, p < .0001$ ) and 81% longer in the (avatar-control, free-space) condition ( $F(1,37) = 10.51, p < .005$ ). The longer average durations observed for the 1.85 m target were not statistically significant (resp. 9% ( $F(1,42) < 1$ ), and 22% ( $F(1,38) = 2.96, p = .09$ )). We are thus able to confirm that



**Fig. 11.** Immersive display exploited for the three strategies: (a) same-height  $H$  strategy, (b) visuocentric strategy for controlling a smaller avatar  $H'$ , (c) egocentric strategy for controlling the same smaller avatar through a  $1/R$  scaling of the virtual environment (visible through the virtual environment height  $W'$  and the fact that the avatar has the same size as the subject).

appears to be relatively more difficult to reach with the avatar control (the curve low height extremum is more than 50% above the minimum in Fig. 13b compared to 30% above the minimum in Fig. 13a). In both cases the minimum lies between 0.6 and 0.7 body height. This is slightly less than in Fig. 2c but the experimental set-up is different in terms of distance to the target in the horizontal plane.

#### 6.4. Scaling strategy factor

Based on prior observations, absolute durations could slightly increase for the highest target as the average of the differing-heights is slightly smaller (1.72 m) than the one of the participants' height (1.76 m).

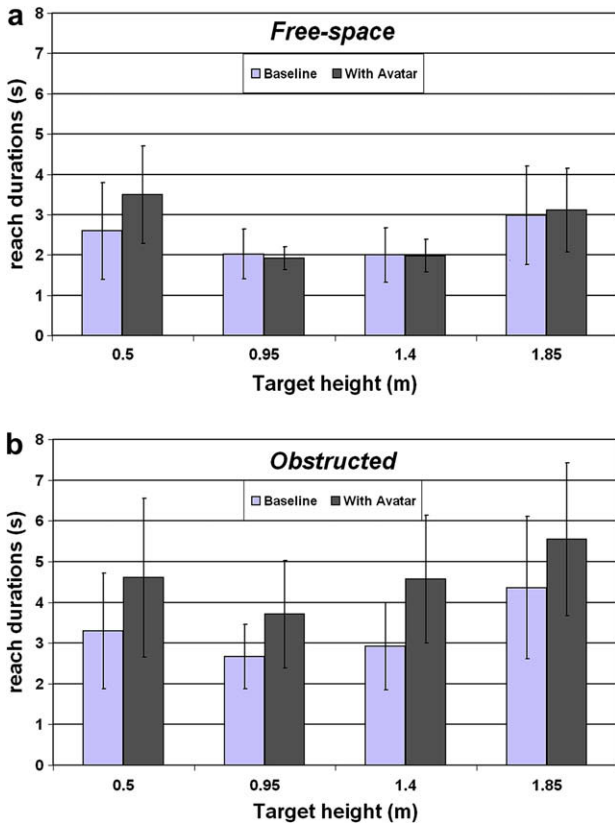
##### 6.4.1. Comparison of absolute reaching durations

Let us first compare the top row of Fig. 14 (differing-height avatar) with Fig. 12a (same-height avatar) for the free-space reach. No

significant difference appears among the baseline data; they are very consistent both in amplitude and variance across the reference and the two scaling strategies.

On the other hand the avatar control requires around 20–30% longer average durations for both scaling strategies compared to the same-height avatar control. However, this is significant only for one of them, namely the 0.95 m target height for which the visuocentric average duration is 30% longer than the same-height strategy ( $F(1,20) = 9.48$ ,  $p = .006$ ). The bottom row of Figs. 14 and 12b treat the data of obstructed reaches. Both the baseline and the avatar control show consistent behaviours between the same-height strategy and the scaling strategies. Although some small differences exist within the baseline control, no significant difference can be detected due to the larger variances. Likewise, no significant difference appears for the avatar control.

Regarding the second hypothesis comparing visuocentric and egocentric strategies, the free-space averages of the former ap-



**Fig. 12.** Average and standard deviation of the absolute reach durations for a same-height controlled entity; (a) free-space, (b) obstructed, (light columns) baseline, (dark columns) avatar.

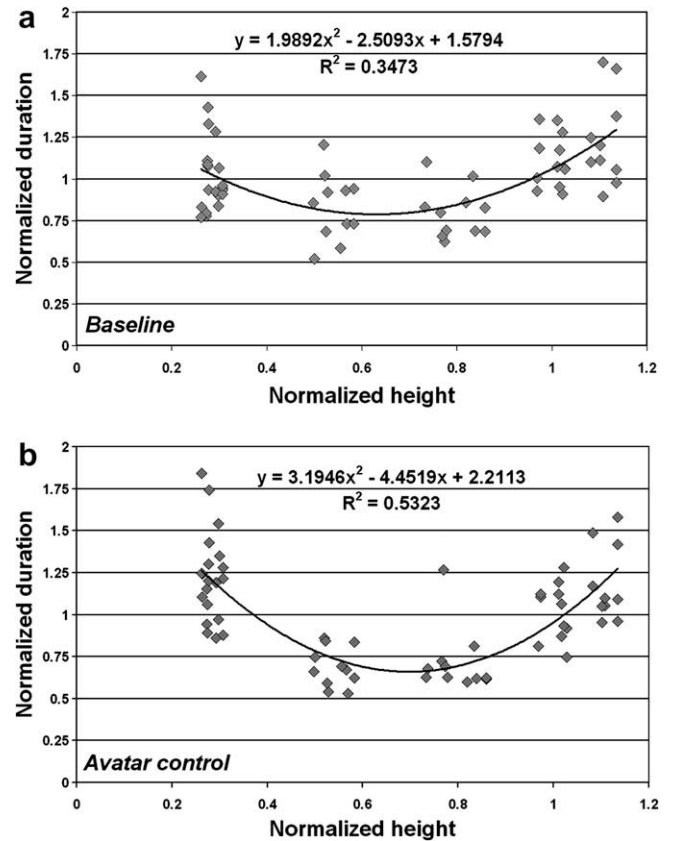
pear to be slightly larger than the one obtained for the latter. However, this is significant only for the 0.5 m target height for which it is 40% longer ( $F(1,40) = 7.03, p = .011$ ). Likewise, no significant differences appear for the obstructed level. Nevertheless it should be noted that the variances are larger for the visuocentric modality.

**6.4.2. Comparison of normalized reaching characteristics**

The main difference in the building process of these characteristics comes from the fact that the eleven subjects have controlled only three distinct height-differing entities, hence the three clearly visible columns of data points per target height on the following illustrations.

First, in the baseline context, the normalized characteristic is very similar for the two scaling strategies (Fig. 15a, Fig. 16a) as for the reference strategy (Fig. 13a). The values of the quality of fit lie in the same range: 0.31 for the visuocentric and 0.43 for the egocentric scaling. The minimum location and the slight asymmetry towards greater fitted values for the highest target are consistent with the reference strategy. This indicates that – when used with the minimal hand proxy display – both scaling strategies succeed in allowing a given subject to replicate the reaching behaviour of a height-differing individual.

On the other hand, the avatar control results in more dissimilar characteristics as can be seen in Figs. 15b and 16b. First, the greater variance compared to the reference case (Fig. 13b) is confirmed by the lower values of the quality of fit: 0.42 for the visuocentric and 0.33 for the egocentric scaling. Although the egocentric scaling characteristic (Fig. 16b) appears to be flatter, similar to its baseline characteristics (Fig. 16a), the larger variance reduces its predictive power.



**Fig. 13.** Normalized reach durations for the reference strategy (A) in free-space as a function of normalized height for a same-height controlled entity; (a) baseline, (b) avatar.

**7. Discussion and conclusion**

**7.1. Viewpoint type**

The study clearly confirms the prior findings that the use of a first-person viewpoint together with an HMD leads to significantly longer reaching response durations for the low target (Section 6.2). The primary cause is the limited field of view. This is confirmed by the fact that targets which were in the field of view from the start did not induce significant differences. In addition, subjects' feedback also included some comfort issues about using the HMD. Even though new HMD devices are now proposed with larger FOV, we suspect that their weight might still impede performance for tasks involving full-body movements and head reorientation.

**7.2. Overview of the results obtained with the baseline modality**

First, the baseline measurements are consistent with the prior findings on full-body reaching described in (Boulic et al., 2006) and illustrated on Fig. 2. The baseline absolute reaching durations in free-space appear to be stable across the two scaling strategies (no significant difference); the same stability of reaching durations can be reported for the obstructed reach difficulty level but with greater variances for all target heights. These variances prevent from characterizing the visibly lower average durations obtained with the egocentric scaling compared to the visuocentric scaling. However the normalized reaching characteristics offer a much better quality of fit for the egocentric scaling (0.42 vs 0.31). Therefore we advocate for retaining the egocentric scaling when impersonating a differing-height mannequin with the baseline modality. The

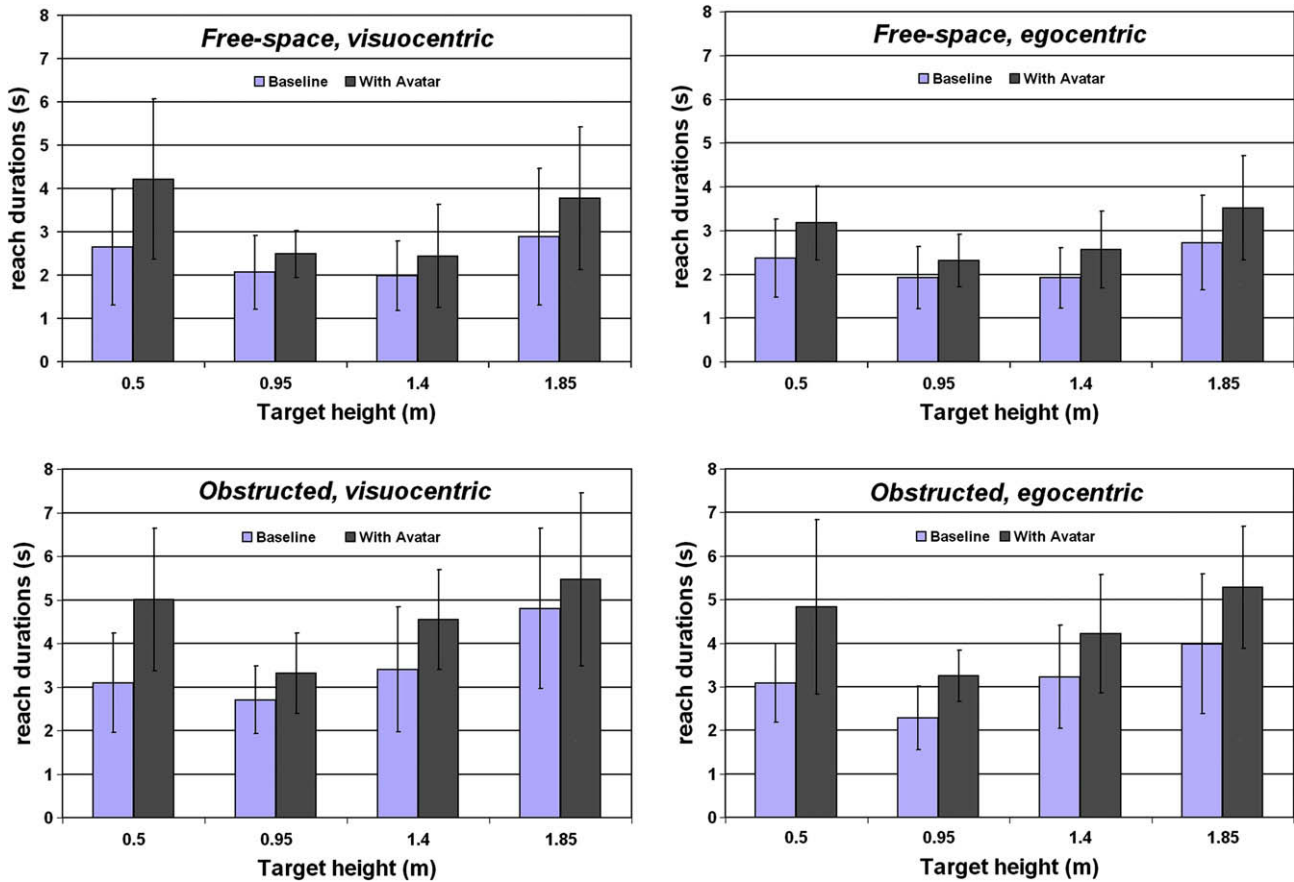


Fig. 14. Average and standard deviation of the absolute reaching durations for a height-differing controlled entity; (top row) free-space, (bottom row) obstructed, (left) visuocentric strategy B, (right) egocentric strategy C, (light columns) baseline, (dark columns) avatar.

next sections summarize the comparisons obtained with the full-body avatar modality.

### 7.3. Same-height avatar control

As discussed in the introduction, the baseline modality is often not sufficient to evaluate interactions with complex virtual environments. Thus we now examine whether the full-body avatar control differs from the baseline modality when controlling a same-height entity.

It is first interesting to point out that the free-space reach bears no significant difference between the baseline and the avatar control, except for the lowest target that produces a significantly longer response. This can be explained by the increased difficulty induced by the numeric IK of the avatar as it has to flex the leg and/or bend the torso to reach the lowest target. The insufficient maturity of the software layer in charge of converting the sensor input into an updated avatar posture may introduce a small lag for low targets. We predict that, with suitable software technology, no difference between the baseline and the avatar control should emerge in free-space reach.

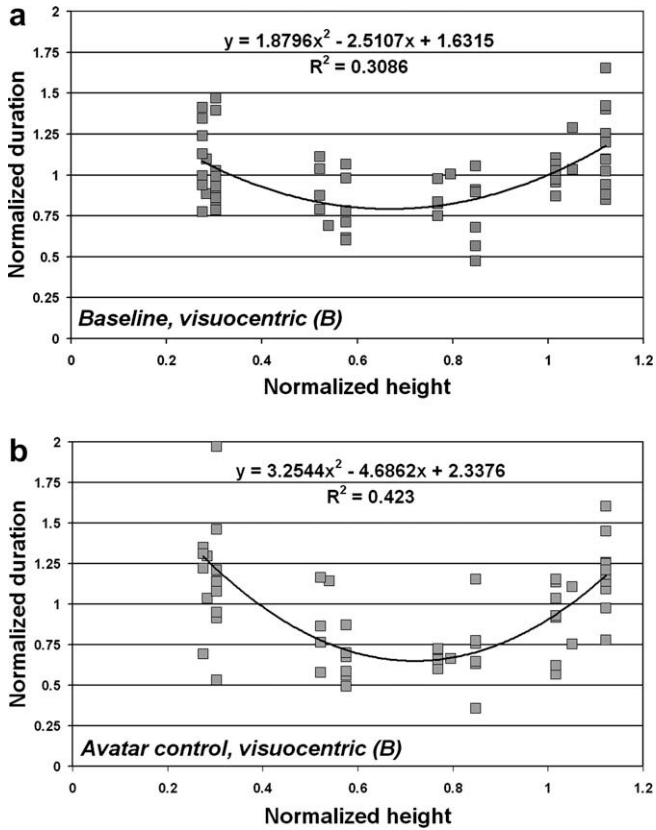
On the other hand, the obstructed difficulty level leads to significantly longer responses of the avatar control compared to the baseline modality, for all target heights. We explain this marked difference by the increased cognitive load due to the additional task of adjusting one's posture so that the avatar posture does not collide with a virtual obstacle. As a side note, such an adjustment requires a third-person viewpoint to view the full avatar body; the location of the first-person viewpoint and its limited FOV would make such a posture evaluation task rather difficult.

With such clear differences between the baseline and the avatar control, we recommend to evaluate virtual prototypes with both modalities, the baseline producing a more relevant duration that is subsequently validated by activating the avatar modality to check for potential collisions. The next section examines the influence of the scaling strategies when controlling a differing-height avatar.

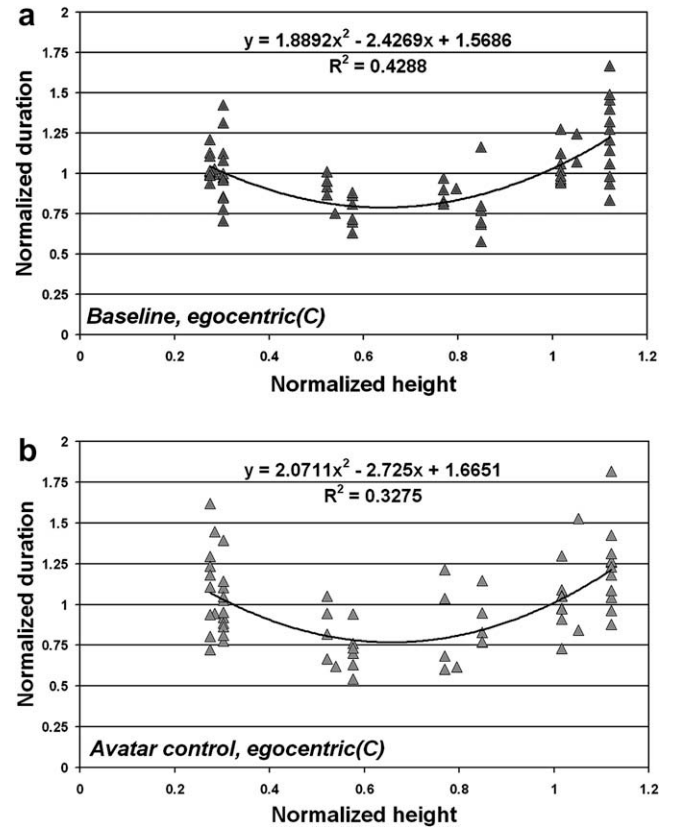
### 7.4. Differing-height avatar control

First, the absolute reaching durations highlight some longer duration when controlling the avatar in free-space for the scaling strategies but few appear to be significant (Section 6.4.1). This could partly be due to the slightly lower average of the differing-heights distribution compared to the participants' height distribution. To better assess that the scaling strategies are achieving their intended goal, it is more relevant to examine the performances on a group basis, i.e. one group including all participants controlling a taller avatar than them (noted taller avatar group), and a second group including all participants controlling a smaller avatar than them (resp. smaller avatar group). We observed that both scaling strategies consistently resulted in reduced durations compared to the same-height strategy when the taller avatar group was reaching the highest target (resp. longer durations for the smaller avatar group). These findings support that the scaling strategies succeed in allowing a participant to replicate the reaching behaviour of a differing-height individual.

Regarding the detection of potential differences among the two scaling strategies, we found only one significantly longer duration for reaching the lowest target in free-space with the visuocentric



**Fig. 15.** Normalized reaching durations for visuocentric scaling (B) in free-space as a function of normalized height for a height-differing controlled entity; (a) baseline, (b) avatar.



**Fig. 16.** Normalized reach durations for egocentric scaling (C) in free-space as a function of normalized height for a height-differing controlled entity; (a) baseline, (b) avatar.

strategy compared to the egocentric strategy. It is interesting to notice that the distance to this target is the greatest among all targets and that it is even amplified for the taller avatar group. The ecological advantage of the egocentric scaling might be more visible for even more distant targets and/or greater differing-heights. A refined experimental protocol would be necessary to generalize this finding.

7.5. Conclusion

We summarize here the recommendations and future research directions we propose for the full-body postural control of virtual mannequins in complex virtual environments.

The findings of this study first confirm that a third-person viewpoint is more appropriate than a first-person viewpoint when reaching performance is to be evaluated; this is due to the HMD limited field of view and additional comfort issues. The first-person viewpoint being potentially useful for evaluating what the avatars sees, we suggest to pursue research on using large immersive screens or CAVE that would allow to alternate transparently between third and first-person viewpoints depending on the current needs.

Second, we recommend exploiting both the baseline (i.e. displaying only a proxy) and the full-body avatar modalities depending on the evaluated feature. The baseline modality is best suited for evaluating reach task duration while the full-body avatar is necessary for collision checking and posture validation purposes. Regarding the effectiveness of the avatar control, a more responsive inverse kinematics technology needs to be developed to reduce the difference observed between the baseline and the avatar control in free-space reaching. Another direc-

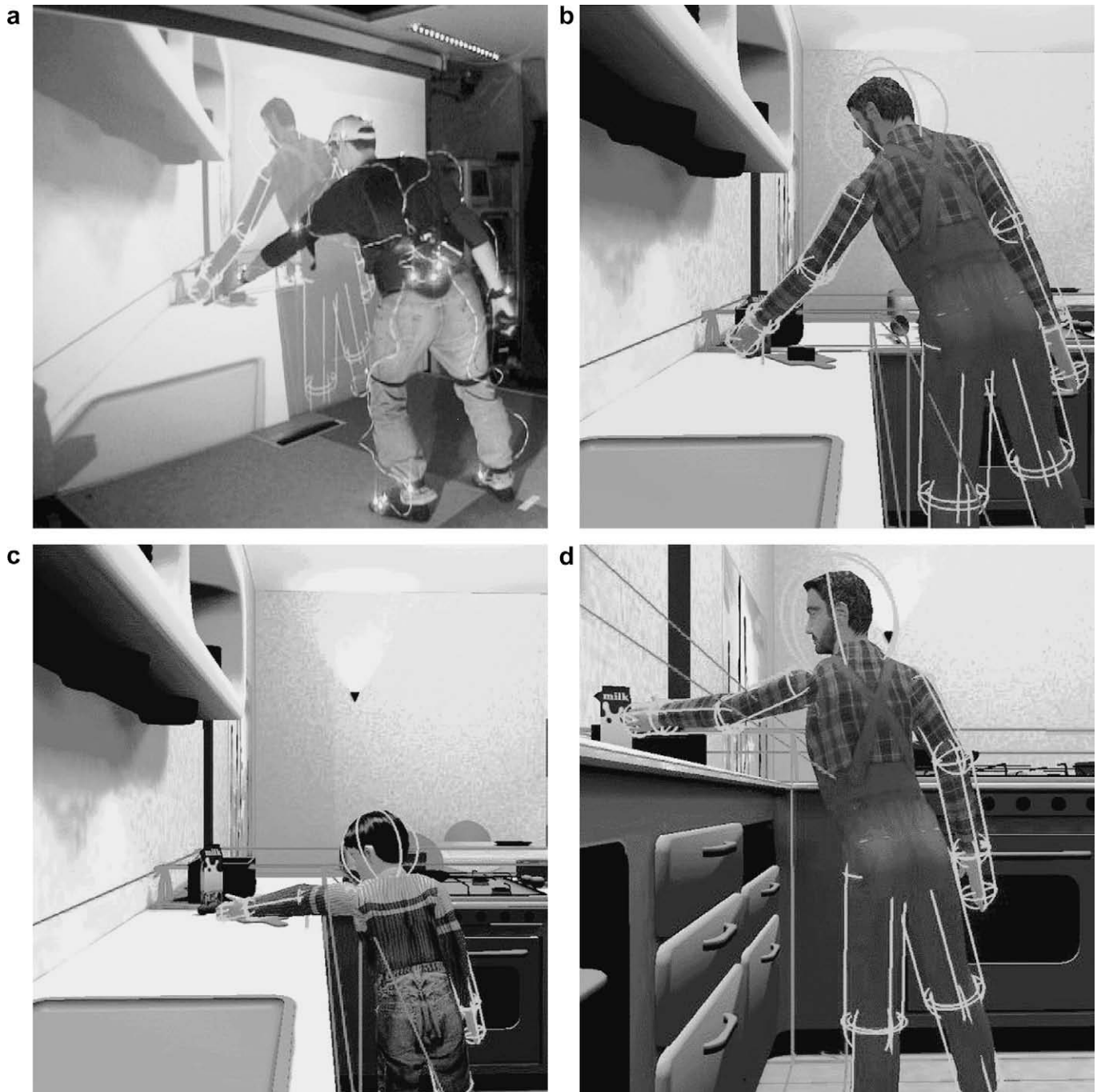
tion of research is to study the minimal feedback (visual or touch sensors) that could be added to the baseline modality to inform the user about on-going collisions of the full-body with the virtual environment.

Third, the experiment has shown the effectiveness of the scaling strategies for the control of height-differing avatars. Although significant statistical evidence is lacking, we would recommend to use the egocentric scaling strategy as it is more ecologically relevant for the person performing the full-body movements. We suggest to conduct a complementary study enlarging the range of height differences to confirm this suggestion.

Based on these conclusions we are confident that we can leverage on the intrinsic enactive knowledge of full-body movements for evaluating complex virtual prototypes. We are currently exploring the full-body control of an avatar interacting in a cluttered virtual environment; we focus on automatically preventing collisions of the avatar (Peinado et al., to appear) to reduce the user cognitive load. Fig. 17 illustrates this problematics combined with the exploitation of the scaling strategies for controlling a child avatar posture.

Acknowledgements

The authors would like to thank the reviewers for numerous constructive remarks, Helena Grillon and Schubert Ribeiro de Carvalho for proofreading, Daniel Raunhardt for his support of the Inverse Kinematics library (Swiss National Science Foundation under the grant 200020-109989) and Achille Peternier for the MVISIO graphic engine with humanoid rendering. We are particularly grateful to Mireille Clavier from the VRLAB, and to Mustafa Kasap and Nedjma Cadi from MIRAlab for pro-



**Fig. 17.** Illustration of the three avatar control strategies in a kitchen environment with a child avatar: (a) immersive display, (b) display of the reference strategy, (c) display of the sensor scaling strategy to control the child avatar, (d) display of the inversely scaled environment for strategy C.

viding the anthropometric avatar models. This research was partially supported by the European network of Excellence ENACTIVE.

## References

- Autodesk MotionBuilder. Available from: <<http://usa.autodesk.com>>.
- Badler, N., Hollick, M.J., Granieri, J.P., 1993. Real-time control of a virtual human using minimal sensors. *Presence Teleoperators and Virtual Environments* 2 (1), 82–86.
- Baerlocher, P., Boulic, R., 2004. An inverse kinematic architecture enforcing an arbitrary number of strict priority levels. *The Visual Computer* 20 (6), 402–417.
- Boulic, R., Maupu, D., Thalmann, D., 2006. Considering the normalized vertical reach performance for consistently controlling virtual mannequins from full-body input. In: *Proc. of the Third International Conference on Enactive Knowledge, ENACTIVE06, Montpellier, 20–21 November 2006*, pp. 51–52.
- Bowman, D., Hodges, L.F., 1997. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In: *Proc. of the 1997 Symposium on Interactive 3D Graphics*, pp. 35–38.
- Caird, J.K., 1994. The effect of virtual hand size on the perception of grasp extent. *Ergonomics and Design* 4, 403–405.
- Carello, C., Groszofsky, A., Reichel, F.D., Solomon, H.Y., Turvey, M.T., 1989. Visually perceiving what is reachable. *Ecological Psychology* 1, 27–54.
- Draper, M., 1995. Exploring the influence of a virtual body on spatial awareness. Master's Thesis. Human Interface Technology Laboratory, University of Washington, Department of Engineering.
- Gabbard, C., Cordova, A., Lee, S., 2007. Examining the effects of postural constraints on estimating reach. *Journal of Motor Behavior* 39 (4), 242–246.
- ISPR, International Society for Presence Research. Available from: <<http://www.temple.edu/ispr/>>.

- Kasap, M., Magnenat-Thalmann, N., 2007. Parameterized human body model for real-time applications. In: Proc. of IEEE International Conference Cyberworlds 2007, 24–26 October 2007, pp. 160–167.
- Kraal, J.C., Baron, E.S., Arbitter, D.S., 2000. Digital occupant: personal immersion for subjective evaluations of a vehicle. In: Proc. of SAE Conference on Digital Human Modeling for Design and Engineering, Dearborn, MI, USA, June 2000. Document number 2000-01-2154.
- Loomis, J.M., Knapp, J.M., 2003. Visual perception of egocentric distance in real and virtual environments. In: Hettinger, I.L.J., Haas, M.W. (Eds.), *Virtual and Adaptive Environments*. Routledge, Taylor & Francis, pp. 21–46, ISBN 080583270X, 9780805832709.
- Mantel, B., Bardy, B.G., Stoffregen, T.A., 2005. Intermodal specification of egocentric distance in a target reaching task. *Studies in Perception and Action* VIII, 173–176.
- Maupu, D., Boulic, R., Thalmann, D., 2007. On-line avatar control using prioritized inverse kinematics. In: Proc. of Fourth International Conference on Enactive Knowledge, ENACTIVE07, Grenoble, 19–22 November 2007, pp. 185–188.
- Menache, A., 1999. *Understanding Motion Capture for Computer Animation and Video Games*. Morgan Kaufmann.
- Mine, M.R., Brooks, F.P., Sequin, C.H., 1997. Moving objects in space: exploiting proprioception in virtual-environment interaction. *Computer Graphics* 31, 19–26.
- Molet, T., Boulic, R., Rezzonico, S., Thalmann, D., 1999. An architecture for immersive evaluation of complex human tasks. *IEEE Transaction in Robotics and Automation, Special Section on Virtual Reality* 15 (3), 475–485.
- Pausch, R., Shackelford, M.A., Proffitt, D., 1993. A user study comparing head-mounted and stationary displays. In: *Proceedings '93 IEEE Symposium on Research Frontiers in Virtual Reality*, pp. 41–45.
- Peinado, M., Meziat, D., Maupu, D., Raunhardt, D., Thalmann, D., Boulic, R., to appear. Full-body avatar control with environment awareness. In: *IEEE CGA*, vol. 28(6).
- Peternier, A., Vexo, F., Thalmann, D., 2008. The mental vision framework: a platform for teaching, practicing and researching with computer graphics and virtual reality. In: *LNCS 5080 Transactions on Edutainment I*, pp. 242–260. Phasespace. Available from: <<http://www.phasespace.com/>>.
- Schafer, W.A., Bowman, D.A., 2004. Evaluating the effects of frame of reference on spatial collaboration using desktop collaborative virtual environments. *Virtual Reality* 7, 164–174.
- Scheidt, R.A., Conditt, M.A., Secco, E.L., Mussa-Ivaldi, F.A., 2005. Interaction of visual and proprioceptive feedback during adaptation of human reaching movements. *Journal of Neurophysiology* 93, 3200–3213.
- Shin, H.J., Lee, J., Shin, S.Y., Gleicher, M., 2001. Computer puppetry: an importance-based approach. *ACM Transactions on Graphics* 20 (2), 67–94.
- Stoakley, R., Conway, M., Pausch, R., 1995. Virtual reality on a WIM: interactive worlds in miniature. In: Proc. of CHI, pp. 265–272.
- Sturman, D.J., 1998. Computer puppetry. *IEEE Computer Graphics and Applications* 18 (1), 38–45.
- Ware, C., Jessome, D.R., 1988. Using the bat: a six-dimensional mouse for object placement. *IEEE Computer Graphics and Applications* 8 (6), 65–70.
- Ware, C., 1990. Using hand position for virtual object placement. *The Visual Computer* 6 (5), 245–253.