Vis Comput (2008) 24: 923–931
DOI 10.1007/s00371-008-0289-x

ORIGINAL ARTICLE

# Visual–haptic perception of compliant objects in artificially generated environments

**Martin Kuschel · Franziska Freyberger ·**
**Berthold Färber · Martin Buss**

**Abstract** Perception of compliant objects through a human system interface with visual–haptic feedback was investigated. Participants had to explore virtual cubes at different compliances by squeezing them with their fingers and observing them visually and haptically. The cubes were rendered by admittance control. Perception of compliance was analyzed using an adaptive staircase method. Results showed that visual–haptic perception of compliant environments is less accurate than perception of position and force stimuli. Furthermore, due to the important role of the visual feedback cross-modal comparisons are more difficult than bimodal comparisons.

**Keywords** Visual–haptic perception · Human system interface · Virtual reality · Telepresence

## 1 Introduction

A human system interface (HSI) enables a human operator to perceive and act in virtual or remote environments (see Fig. 1). Perception and manipulation capabilities can be increased by reflecting multimodal (e.g. visual and haptic) information of the target environment (for an overview, see on force reflecting presence systems [2]). Kinesthetic haptic (by now *haptic*) feedback is sensitive since mechanical energy is exchanged over command and feedback signals. This closed feedback loop is susceptible to different kinds of disturbances and can even become instable (e.g. [19, 42]). Stability measures can deteriorate display performance and, consequently, at the operator's site incongruences between visual and haptic information can occur. However, humans can perceive even incongruent bimodal information without any conflict. Measurement of this perceptional process requires an HSI with high accuracy and extensive experiments using psychophysical procedures. Perception of bimodal mechanical information is analyzed here at the example of compliance information.

It is known that information of more than one modality is integrated to a coherent percept (e.g. [33, 36]). Precondition of integration is spatial as well as temporal congruence of information (e.g. [33, 36]). However, if an intermodal conflict is below threshold of perception, integration still takes place (e.g. [4]), even though only one modality has been attended (e.g. [23]). In this context, Marks introduces the terms *stimulus* and *perceptual congruence* (see [4], pp. 85–105): Stimulus congruence denotes that there are no differences in the physical stimulus parameters. Perceptual congruence is a psychological construct: Even without physical congruence discrepancy in visual–haptic information remains to some extent unnoticed by the observer. Because a qualitative difference of the integrated percept depending on physical or perceptual congruent stimulus presentation has been observed, different mechanisms may be involved [22, 32]. Attentional processes might influence integration and the extent of permissible discrepancy, respectively (see

M. Kuschel (✉) · M. Buss
Institute of Automatic Control Engineering (LSR), Technische Universität München, Munich, Germany
e-mail: martin.kuschel@tum.de

M. Buss
e-mail: mb@tum.de

F. Freyberger · B. Färber
Human Factors Institute (IfA), University of the Bundeswehr Munich, Munich, Germany

F. Freyberger
e-mail: franziska.freyberger@unibw.de

B. Färber
e-mail: berthold.faerber@unibw.de

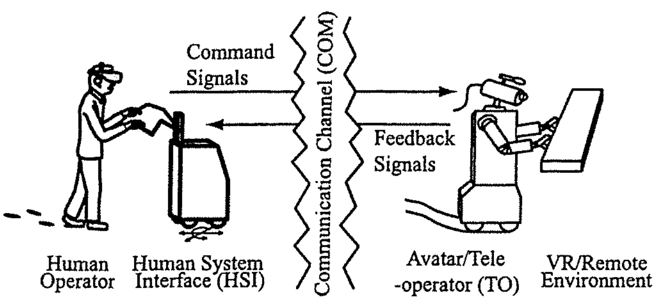


**Fig. 1** Multimodal VR/Telepresence: Visual and haptic information is reflected increasing the human operator's performance in a target environment

e.g. [10, 33]). Research on multimodal integration encountered that information of two or more modalities is fused by differently weighting it (e.g. [5, 31]). The relative contribution of each sense depends either on the appropriateness (e.g. [39]), the effectiveness (e.g. [3]) or the reliability (e.g. [6, 7]) of each modality, or on the direction of focusing attention (e.g. [13]). A wealth of research into visual–haptic integration exists and has concentrated either on size (e.g. [6, 18, 27]), shape (e.g. [15, 16, 28]) or texture perception (e.g. [12, 14, 20]), as well as on visual influence on proprioceptive localization (e.g. [25, 26, 38]), respectively. Most of the research on integration indicates visual dominance over the haptic modality (for a review, see [4, 39]), especially in spatial properties (e.g. [3]). Moreover, some factors, such as age (e.g. [24]), response modality (e.g. [17]), instruction (e.g. [21]) or noise (e.g. [6]), have been found to reduce visual influence. Some studies even show tactile dominance over vision (e.g. [8, 12, 20]) or at least in some tasks (e.g. [9]). Visual dominance therefore is no general phenomenon and depends on additional task relevant factors.

The contribution of this article is to identify the just-noticeable difference (JND) (see e.g. [11]) when perceiving object compliance through a visual–haptic HSI. The JND is defined as

$$\mathrm{JND} = \frac{|S - S_{ref}|}{S_{ref}}. \qquad (1)$$

Compliance is the combination of force $f$ and position $x$ information and can be expressed by Hooke's law

$$S = \frac{x}{f}. \qquad (2)$$

Bimodal perception of compliance information has evoked only few studies yet [35, 41]. Three hypotheses were tested.

**Hypothesis 1** Attending separately to information of two modalities arising from one source might introduce a high bias: The human perceptual system tries to integrate even conflicting information to provide a coherent percept, especially if information is derived from one source (see [33, 36]). Therefore, concurrent comparison of visual and haptic information should result in reduced detection performance (further referred to as 'method A'). On the other hand, a low discrepancy threshold should result when attending to the object as a whole, and hence to compare a visually presented object to an haptically presented object sequentially (further referred to as 'method B').

**Hypothesis 2** Most of the research on visual–haptic integration reports visual capture in intermodal conflict situations (see above). It is therefore expected that visual dominance should occur: The detection of intermodal discrepancies in object compliance should be impaired when the visual modality remains unaltered (as the target or reference modality) and hence the haptic modality varies.

**Hypothesis 3** As has been shown in different studies (see above), visual dominance seems not to be a general phenomenon and to depend on task-relevant factors. It is expected that modality dominance is not constant over the whole stimulus range of object compliance: Low compliant objects provide scant visual compliance information, whereas high compliant objects might be easier perceived when relying on visual information.

All three hypotheses were tested regarding influence of assessment method, reference modality (i.e. unaltered target modality), and reference compliance with respect to the detection threshold.

In Sect. 2 the HSI is explained and in Sect. 3 the experimental assessment is described. The results are presented in Sect. 4 and discussed in Sect. 5.

## 2 Visual–haptic compliance rendering

### 2.1 Hardware and software

Haptic information is exchanged via a haptic interface comprised of two self-made SCARA robots providing a single degree of freedom each. The system interacts with index finger and thumb to allow gripping movements. High fidelity components like *Maxon* motors and *Harmonic Drive* gears enable best possible control. Workspace is about 80 mm and maximum force is about 50 N. Position information is measured by angle encoders and force is sensed by strain gauges attached on both robot links. Visual information is provided over a TFT screen. Thereby, the compliant environment is represented by a gray cube squeezed by two orange spheres (on opposed cube sides) representing finger positions. See Fig. 2 for an illustration of the overall system and Fig. 3 for

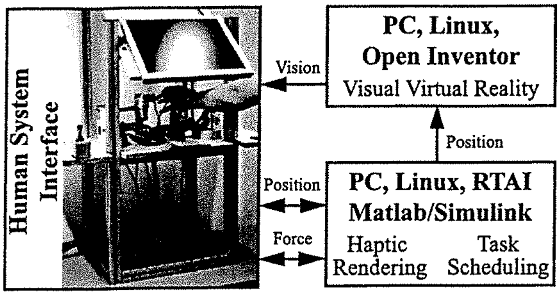


**Fig. 2** Human system interface and real-time processing unit: Visual and haptic information is exchanged and positions and forces are measured

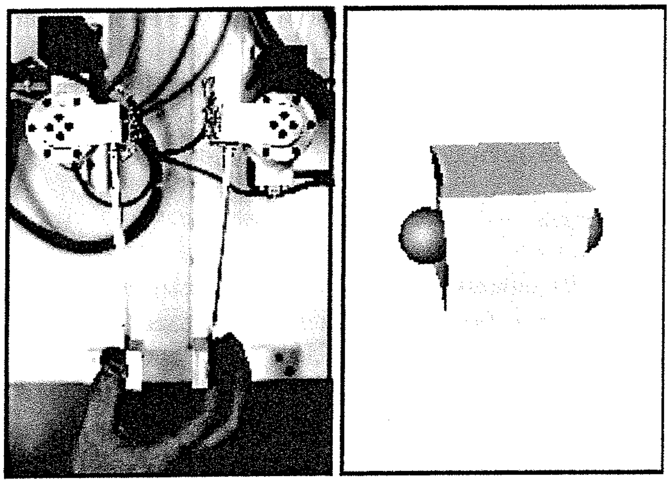


**Fig. 3** Haptic and visual feedback: The haptic feedback renders a compliant cube to be explored by thumb and index finger. In the visual feedback fingers are replaced by orange spheres

close-ups on the two types of feedback. The TFT screen is slanted by 40° and mounted in the line of sight to the hand enabling participants to look at the display as if they were looking at their hand.[1]

The system is connected to a PC running *RTAI*-RealTime Application Interface for Linux. SCARA sensor signals are recorded by a *Sensoray 626* DAQ-Card providing 16-bit sensing resolution. Signal processing algorithms are implemented as *Matlab/Simulink* models with real-time code generated automatically. The system operates at 1 kHz sampling frequency. Measured positions are transferred to a second PC running the visual VR programmed in *Open Inventor*.

## 2.2 Kinematics and dynamics

The identical robots of the HSI are controlled independently using the same admittance control scheme (see Fig. 4 for kinematical configuration). In the following, the concept is explained using a single-robot system without loss of generality. Kinematical transformations' (forward kinematics, inverse kinematics) mapping torques $T$ to forces $f$ are omitted for simplicity.

For dynamics consider a mechanical robot with a single translational degree-of-freedom. The dynamical equation is given by

$$M_h\ddot{x}_h + D_h\dot{x}_h + K_h x_h + n_h = g_h - f_o^e, \tag{3}$$

where $M_h$, $D_h$, $K_h$ denote mass, damping, and stiffness of the HSI and $n_h \in \mathbb{R}$ the nonlinear dynamics of the HSI. Robot force $g_h \in \mathbb{R}$ depends on motor torque $T$ and on link length $l$. The position of the end effector is denoted by $x_h$. Input–output linearization [30] is achieved by commanding

$$g_h = f_h^m + n_h. \tag{4}$$

The resulting linear dynamics are

$$M_h\ddot{x}_h + D_h\dot{x}_h + K_h x_h = f_h^m - f_o^e, \tag{5}$$

where $f_h^m$ is the new motor force of the linearized HSI.

A PID controller, $C : U \to M$, realizes the control signal $f_h^m$ according to the position difference of robot and stimulus compliance

$$f_h^m = C[x_s - x_h], \tag{6}$$

where the brackets indicate that $C$ contains differential and integral operations. The HSI is connected to the human operator. The velocity of the HSI and the velocity of the operator's fingers are opposite, hence

$$x_h = -x_o. \tag{7}$$

The dynamics of the robot interacting actively with the human operator is described by

$$f_o = M_h\ddot{x}_o + K_o x_h + f_o^m, \tag{8}$$

where $M_o$, $K_o$ denote mass and flexibility of the operator's fingers and $f_o^m$ is the force actively intended by the human operator impeded by the force $f_o$ that mediates the stimulus compliance (VR displayed to the operator).

The dynamics of the stimulus compliance is described by the admittance $S : U \to M$ (force input, position output), which represents compliance according to Hooke's law

$$x_s = S f_o^e, \tag{9}$$

where $S$ [mm/N] is the compliance whose perception is addressed in this study. The control concept employing inner position control driven by a VR with force reference is called admittance control. It is best suitable for rendering non-rigid environments like compliant environments (see [37] for detailed information). Minimum compliance (= maximum stiffness) that can be rendered is $S = 0.2$ mm/N.

---

[1] The tool transformation has no influence on the dynamics of the gripping movement, if participants are given a learning phase (e.g. see [1]).

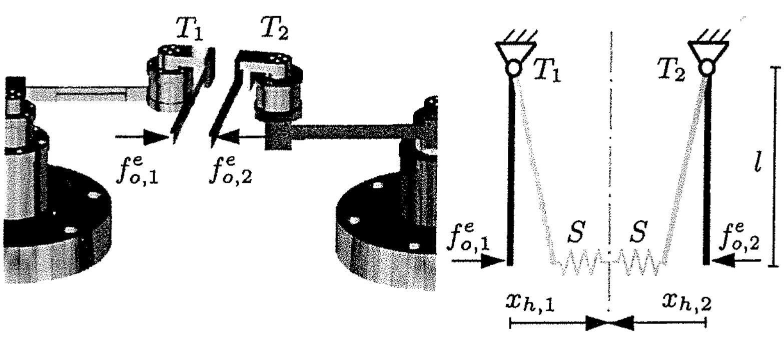


**Fig. 4** Kinematical structure of the haptic display: Two SCARA robots haptically render compliant cubes for gripping movements

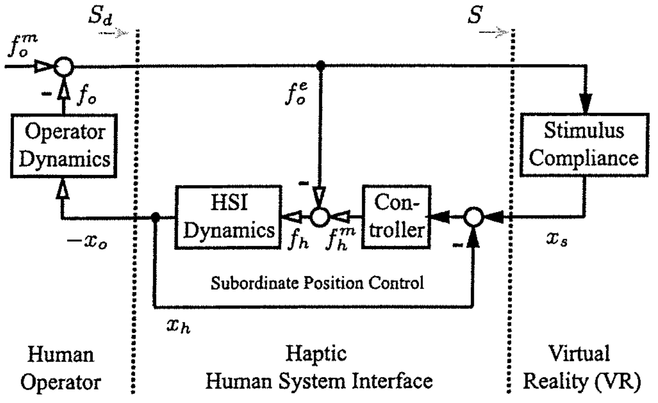


**Fig. 5** Control of HSI: Different compliances are rendered using a high-fidelity robot driven by admittance control. The displayed stimuli compliances $S_d$ show nearly no differences to the commanded stimuli compliances $S$. (*Hollow arrows* represent physical interaction, *filled arrows* represent signal processing)

A block diagram of the human operator interacting with the admittance-controlled HSI is depicted in Fig. 5. Hollow arrows depict physical interactions, filled arrows are used for signal interactions. All subsystems are considered to be linear(ized) and time-invariant. The fidelity of the VR depends on dynamics and control of the HSI. The robot is light-weighted, dynamics of the motor current control are negligible, and velocities are small (i.e. friction effects negligible). Consequently, the transparency of the system can be assumed nearly ideal and the displayed dynamics $S_d$ can be considered equal to the dynamics of the VR

$$S_d = S. \tag{10}$$

## 3 Experimental assessment

### 3.1 Participants

Thirty-two (32) students of the Technische Universität München and the Universität der Bundeswehr München took part in this study and were paid for participation. Half of the participants were assigned to group A (method A), the other half to group B (method B). Due to missing values, two (group A) and five (group B) participants had to be excluded from further analysis. The average age of participants amounted to 25 years (group A) and 26 years (group B). Eleven (11) men and 3 women (group A) and 6 men and 5 women (group B) participated. All of them were right-handed and had normal or corrected-to-normal vision.

### 3.2 Stimuli

Seven (7) reference compliances were selected covering a broad range feasible by the HSI. Reference compliance amounts to

$$S = [0.2; 0.4; 0.5; 0.8; 1.4; 2.5; 4.9]\,\mathrm{mm/N}. \tag{11}$$

Additionally, intermodal discrepancy should be assessed with either the visual or the haptic modality remaining unchanged and therefore being the reference modality: Reference compliance of the unchanged modality was one of the seven values, whereas compliance of the comparison or non-target modality was varied according to the procedure.

### 3.3 Procedure

An adaptive staircase method was used to assess performance of participants. Thereupon, two ways of assessment were defined: Method A demanded comparing visual and haptic information *concurrently* within one trial and to decide whether sensory information deviates from each other, while in method B participants had to *sequentially* compare a congruent bimodal stimulus with an incongruent bimodal stimulus. With both methods a total of 14 threshold values had to be assessed: each reference stimulus (7) and target modality (2) combination. In order to reduce overall testing time for participants, method of assessment was chosen to be a between-participants variable, and testing of the 14 experimental conditions was divided into two sessions with seven randomly chosen stimuli. These stimuli were selected with the following restrictions: Neither the same reference compliance nor the same reference modality was presented in succession.

Because the ability to be drawn into a book, film or VE, better known as *immersive tendency* (see [40]), has been known to play an important role in designing human–machine interaction, discrepancy of bimodal information might be mediated by this personal factor. Therefore, an additional 12-item questionnaire was included to control for that variable ([40], translated by [29]): Immersive tendency was assessed by answering 12 items building the two factors, *tendency to get emotionally involved* and *degree of involvement* (see [29]).

An additional group-specific question was included in the demographical questionnaire: Group A rated under which reference (unchanged) modality they felt easier to perform the task, whereas group B rated the sensory information they mostly had relied upon during test sessions.

Participants were seated in front of the HSI with their dominant hand grasping the device and while looking nearly perpendicular at the screen. They were carefully instructed according to their group membership to which they were randomly assigned. A training period had to be completed prior to each test session. Afterwards, seven experimental conditions (one test session) were randomly presented. Participants explored the stimulus depending on their group membership and responded by a joystick.

The start and end of each trial was signaled by a sound. Duration of stimulus presentation depended again on group membership: Group A compared visual and haptic information for 4 s with an intertrial-interval amounting to 4 s. Group B tested the stimulus compliance for 2 s with an interstimulus-interval of 2 s and an intertrial-interval of 4 s. Masking of environmental noise was regarded to be not necessary due to the HSI making no disturbing noise, which might influence the participants' responses.

At the end of the second test session participants filled in questionnaires assessing their demographical data, their experience during testing (additional group specific question) as well as the immersive tendency questionnaire.

Two different procedures of measuring the perceptual threshold according to the group variable had been used (method A, B). Both methods assessed the relative just noticeable difference JND. The psychophysical procedure to derive the JND was an adaptive staircase method targeting the 50% performance level. Initial stimulus and step size were adopted according to group membership.

*Group A* Participants were instructed to concurrently compare within one trial information from both modalities given the reference modality which was announced by the experimental instructor prior to measurement. Therefore, the experimental task was to make *cross-modal* comparisons and to choose between two response alternatives, namely 'difference' vs. 'no-difference.' In case of haptic reference the visual modality had to be adjusted to match the haptic modality. Since only position measurement is possible by vision, the perceptual task was to match the visual position to the haptic position encountered while exploring the compliant object

$$\text{Method A, haptic ref.:}\quad x_{vis} \xrightarrow{\text{adjust to}} S = \frac{x_o}{f_o^e}, \tag{12}$$

where $x_{vis}$ denotes the position information obtained by the visual modality and $x_o$ denotes the information measured by the haptic modality as used before. In case of visual reference the haptic modality had to be adjusted to match the visual modality. Since participants had to match compliance information (measured by the haptic modality) to a position (measured by the visual modality). they had to filter out the haptic position $x_o$ from the haptic compliance estimate to perform the matching

$$\text{Method A, visual ref.:}\quad S = \frac{x_o}{f_o^e} \xrightarrow{\text{adjust to}} x_{vis}. \tag{13}$$

Threshold measurement by the staircase procedure started with the comparison stimulus

$$S_0 = 2S_{ref}, \tag{14}$$

while step size was adapted according to

$$S_1 = S_0 - xS_{ref}. \tag{15}$$

After the third turning point, step size was reduced from $x_1 = 0.1$ to $x_2 = 0.03$. Congruent stimuli were interspersed with a probability of 10%. After having reached the threshold ten times, threshold measurement ended: The difference threshold was computed as the mean of the limit cycle consisting of the last, unchanged turning points. The principle of the adaptive staircase method is depicted in Fig. 6.

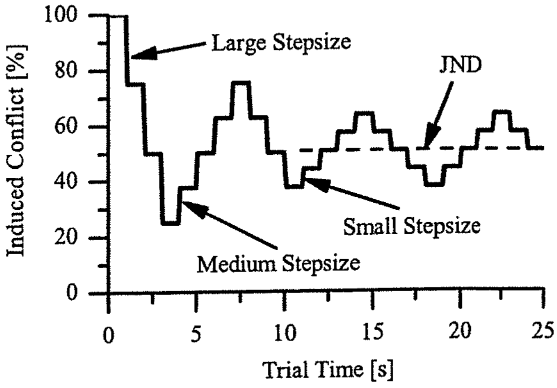


**Fig. 6** Adaptive staircase method for difference thresholds: The *JND* is the mean of the limit cycle of a subject's answers swaying between 'detected' and 'not detected.' The stimulus is adjusted upwards when the stimulus difference (conflict) is not detected, and is adjusted downwards if detected. To save trial time, adjustment steps are large at the beginning and decreased after each lower turning point

*Group B* Participants explored two objects sequentially, one of which characterized by a discrepancy and the other being the congruent one

$$\text{Method B}: \quad S \overset{\text{compare with}}{\longleftrightarrow} \hat{S}, \tag{16}$$

where $S$ denotes the congruently displayed compliance and $\hat{S}$ the incongruently displayed compliance. Reference modality was not announced by the experimental instructor. The participants' task was to decide whether the second stimulus felt more or less compliant than the first or whether there was a difference between the first and second stimulus. Threshold measurement started with the comparison stimulus yielding a discrepancy of

$$S_0 = S_{ref} + 0.8 S_{ref}, \tag{17}$$

while step size was varied according to (15). Until the third transition point has been reached, step size amounted to $x_1 = 0.1$ and was then reduced to $x_2 = 0.03$. Congruent comparison stimuli were interspersed with a probability of 5%. After having reached the threshold six times, the sequence ended. The difference threshold was defined as the mean of the limit cycle consisting of the last, unchanged, turning points.

## 4 Results

### 4.1 Immersive tendency

Participants rated their immersive tendency on a 7-point scale building the two factors, *emotional involvement* and *degree of involvement*, which were computed for each participant. Group A showed an average emotional involvement of 23.3 (standard deviation $sd = 5.8$) and an average degree of involvement of 25.6 ($sd = 6.9$), whereas mean emotional involvement amounted to 16.8 ($sd = 4.4$) and mean degree of involvement to 24.6 ($sd = 8.7$) in group B. All values did not statistically significantly differ from those reported by Scheuchenpflug [29], indicating that the participants are a good sample of population.

In order to find out whether the two groups (A, B) differed from one another regarding the immersive tendency, because there seemed to be a difference in emotional involvement at least descriptively, a t-test for independent groups was computed. No difference in degree of involvement could be found ($t(21) = 0.3$, $p = 0.8$). However, both groups differed statistically significantly in their emotional involvement ($t(21) = 3.0$, $p < 0.05$): Group A rated to be higher emotionally involved than group B. However, only 29% of variance could be explained by this effect, which therefore can be neglected.

*Group A* No correlation between emotional involvement and performance could be observed. Only two variables when the haptic modality was the (unaltered) target modality showed a statistically significant (significance level was 5%) correlation with degree of involvement: Reference compliance of 2.45 mm/N (Spearman $\rho = +0.7$) and of 0.42 mm/N ($\rho = -0.8$). A positive correlation indicates a higher JND along with a higher degree of involvement, whereas a negative correlation indicates a better performance (reduced JND) with a higher degree of involvement.

*Group B* Additionally, emotional involvement had no influence on performance, whereas degree of involvement statistically significantly influenced perception threshold: With the visual modality being the target modality and a reference compliance of 0.42 mm/N, a correlation could be observed ($\rho = -0.7$).

### 4.2 Group specific questions

Participants answered an additional question according to the modality which facilitated the given task (group A) and according to the modality participants mostly attended to (group B). To determine whether there was an influence on performance or a relation to immersive tendency, separate correlation analyses for both groups (A, B) were computed.

*Group A* Participants answered that performing the task was easier when the reference modality was either the haptic ($n = 2$) or the visual ($n = 4$) modality, or both together ($n = 8$). There was neither a correlation with immersive tendency nor performance.

*Group B* Participants reported that they primarily attended the haptic modality ($n = 5$), the visual modality ($n = 3$), both modalities without preference ($n = 2$). Rating of the attended modality significantly affected performance (significance level was 5%) when the visual modality was the (unaltered) target modality: Reference compliance of 0.5 mm/N ($\rho = +0.7$) and 0.42 mm/N ($\rho = +0.8$) indicating higher performance when attending to the haptic modality, medium performance for the visual and lowest performance when attending to both modalities. However, there was no correlation with immersive tendency.

### 4.3 Descriptive analysis

The JND was computed for each experimental condition within both groups. The mean performance is presented in Table 1.

**Table 1** Summary of results: When comparing compliant objects reference modality seems only to play a role if comparisons are performed cross-modal. Sequentially comparing compliances yields better performance and is independent of reference modality

| Reference (target) modality | Comparison Concurrently (group A) | Sequentially (group B) |
|---|---|---|
| Vision | JND = 128% | JND = 55% |
| Haptic | JND = 85% | JND = 55 to 68% |

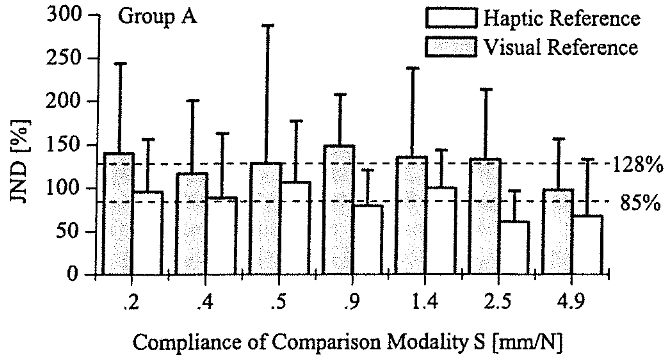


**Fig. 7** Results of Group A: When visual and haptic compliance information of one object were compared concurrently (cross-modally), the JND is higher if the visual modality is the reference modality

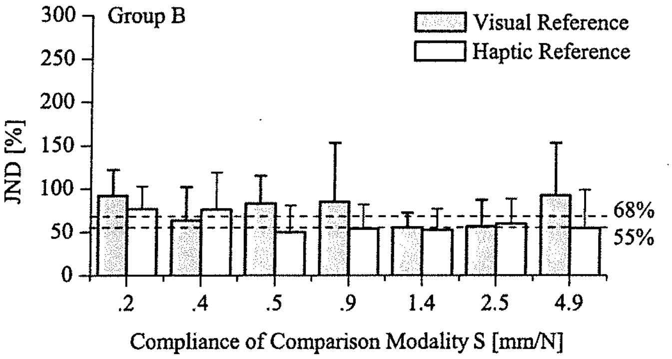


**Fig. 8** Results of Group B: When visual–haptic compliance information of one object was compared sequentially, the JND was higher in only some stimuli when the visual modality is the reference modality. In general, JNDs were much lower compared to the cross-modal task (Fig. 7)

*Group A* As can be seen in Fig. 7, reference modality affects performance: When the visual modality remains unchanged according to (13), the average JND is higher than in the case of the haptic reference according to (12); i.e., the task described by (13) is more difficult than the task described by (12). This indicates visual dominance, i.e. it is more difficult to filter the haptic position information to match the visual reference than to adjust the visual information to a filtered haptic position reference (visual modality measured position only).

*Group B* Group B showed overall higher performance than group A (see Fig. 8): JND is lower and standard deviation is smaller, indicating that group B had less difficulties performing the detection task sequentially according to (16). Again, there seems to be an influence of reference modality on JND, but only in some reference stimuli. Most of the participants who had to be excluded were unable to perform the tasks when the visual modality remained unaltered and thus more likely when attending to the visual modality.

### 4.4 Testing hypotheses

In order to determine the influence of reference compliance (0.22 to 4.88 mm/N) and reference modality (visual, haptic) depending on assessment method (A, B), on the ability to detect discrepancies in intermodal information, a 7 × 2 × 2 analysis of variance (ANOVA) with repeated measurements and method as between-participants variable was computed (significance level of 5%).

Threshold was significantly different in both groups ($F(1,23) = 9.5$, $p < 0.05$; partial $\eta^2 = 0.29$): Group B showed higher sensitivity, i.e. detected smaller discrepancies in intermodal information. Additionally, reference modality influenced the JND: Performance was higher when the haptic modality was reference modality and therefore the visual modality was changed during the testing ($F(1,23) = 8.6$, $p < 0.05$; partial $\eta^2 = 0.27$). No interaction of target modality and reference compliance could be observed ($F(6,138) = 0.7$, $p = 0.6$). No other effects reached statistical significance.

However, effect size (partial $\eta^2$) of both main effects is very low and, as can be seen comparing Fig. 7 to Fig. 8, the above reported influence of reference modality seems primarily due to performance of group A. Therefore, a 7 × 2 ANOVA with repeated measurements was computed for each group.

Again, main effect of modality was statistically significant in group A ($F(1,13) = 10.3$, $p < 0.5$); the effect now accounted for 44% of the variance. However, no influence of reference modality on performance could be observed in group B ($F(1,13) = 1.2$, $p = 0.3$). The only other though negligible effect that reached significance was the interaction between target modality and reference compliance ($F(6,60) = 2.5$, $p < 0.05$; partial $\eta^2 = 0.20$) indicating higher JND when reference compliance is 0.5, 0.85 and 4.88 mm/N and the visual modality is the reference.

## 5 Discussion

Difference thresholds in visual–haptic compliance information were assessed for different experimental conditions. Participants either had to concurrently compare visual with haptic compliance information (method A) or to sequentially compare two compliant objects displayed visually–haptically (method B). The chosen method affected the detection performance, as was expected (*Hypothesis 1*). As can be seen in Table 1, concurrent comparisons yield low performance (around 128% to 85%), whereas sequential comparisons yield performances between 55% and 68%. Similar results are obtained by Srinivasan, Beauregard & Brock who showed that participants' ability to identify the less compliant of two easily distinguishable compliant stimuli decreased as the ratio between visual and haptic discrepant compliance information increased to around 0.5 [35].

The different JNDs reflect that detecting visual–haptic conflicts is very difficult as long as no congruent comparison is available. The perceptual system may integrate the information in order to provide a coherent percept (e.g. [33, 36]). On the other hand, conflicts can be more easily detected when comparing to congruent information. As expected (*Hypothesis 2*), reference modality influenced discrimination performance: When varying haptic information (visual modality is unchanged), the JND is higher than with the haptic modality being the reference (see also [35]). This indicates that participants relied more on visual information when performing the discrimination task. Although this visual dominance can be observed for both groups (main effect), especially performance of group A contributes to this effect since this group had to perform an additional filtering to extract the position information from the haptic percept. In group A, 128% difference between the visual comparison and the haptic reference information is necessary in order to be detected, whereas 85% intermodal difference can be detected when the haptic modality remains unchanged (see Table 1). This indicates that the filtering of position information is more demanding if the information that has to be matched has to be filtered than if the reference information has to be filtered.

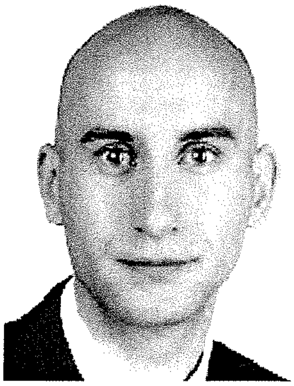
Visual dominance depended on reference compliance in group B, as expected (*Hypothesis 3*). However, effect size is rather low. Additionally, a positive correlation between directing attention and discrepancy threshold could be observed in some reference stimuli: Performance decreased when participants attended to the 'wrong' modality, i.e. vision. Moreover, the analysis of missing values revealed that expecting the 'wrong' modality resulted in a non-convergence of the iterative psychophysical method. The cost of attending the wrong modality has already been shown to decrease performance (e.g. [13, 34]). Whether the influence of attention accounts for the observed result has not been systematically addressed in this study. Further experiments have to clarify, whether this effect might account for this interrelation.

**Acknowledgements** This research was funded by the German National Science Foundation (DFG) within the Collaborative Research Center on "High-Fidelity Telepresence and Teleaction" (SFB 453).

## References

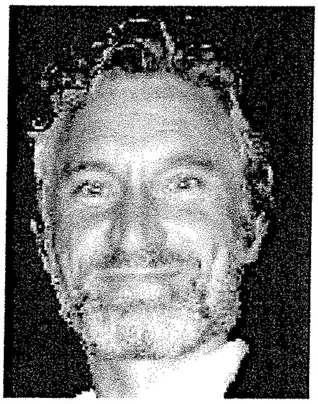
1. Brenner, E., Smeets, J.B.J.: Fast corrections of movements with a computer mouse. Spat. Vis. **16**, 364–376 (2003)
2. Buss, M.: Mechatronics and control issues in multi-modal telepresence. In: IEEE International Conference on Mechatronics and Robotics, MechRob2004 (2004)
3. Calvert, G.A., Brammer, M.J., Iversen, S.D.: Crossmodal identification. Trends Cogn. Sci. **2**(7), 247–253 (1998)
4. Calvert, G.A., Spence, C., Stein, B.E.: The Handbook of Multisensory Processes. MIT Press, Cambridge (2004)
5. Driver, J., Spence, C.: Multisensory perception: Beyond modularity and convergence. Curr. Biol. **10**(13), 731–735 (2000)
6. Ernst, M.O., Banks, M.S.: Humans integrate visual and haptic information in a statistically optimal fashion. Nature **415**, 429–433 (2002)
7. Ernst, M.O., Bülthoff, H.H.: Merging the senses into a robust percept. Trends Cogn. Sci. **8**(4), 162–169 (2004)
8. Ernst, M.O., Banks, M.S., Bülthoff, H.H.: Touch can change visual slant perception. Nat. Neurosci. **3**(1), 69–73 (2000)
9. Fishkin, S.M., Pishkin, V., Stahl, M.L.: Factors involved in visual capture. Percept. Mot. Skills **40**, 427–434 (1975)
10. Garner, W.R.: Attention: The processing of multiple sources of information. In: Handbook of Perception, vol. 2, Psychological Judgement and Measurement, pp. 23–59. Academic, New York (1974)
11. Gescheider, G.A.: Psychophysics—Method and Theory. Wiley, Hillsdale (1976)
12. Guest, S., Spence, C.: Tactile dominance in speeded discrimination of textures. Exp. Brain Res. **150**(2), 201–207 (2003)
13. Guest, S., Spence, C.: What role does multisensory integration play in the visuotactile perception of texture? Int. J. Psychophysiol. **50**, 63–80 (2003)
14. Heller, M.A.: Visual and tactual texture perception: Intersensory cooperation. Percept. Psychophys. **31**(4), 339–344 (1982)
15. Heller, M.A.: Haptic dominance in form perception with blurred vision. Perception **12**, 607–613 (1983)
16. Heller, M.A.: Haptic dominance in form perception: Vision versus proprioception. Perception **21**, 655–660 (1992)
17. Heller, M.A., Calcaterra, J.A., Green, S.L., Brown, L.: Intersensory conflict between vision and touch: The response modality dominates, when precise, attention-riveting judgments are required. Percept. Psychophys. **61**(7), 1384–1396 (1999)
18. Kinney, J.A.S., Luria, S.M.: Conflicting visual and tactual-kinesthetic stimulation. Percept. Psychophys. **8**(3), 189–192 (1970)
19. Lawrence, D.A.: Stability and transparency in bilateral teleoperation. IEEE Trans. Robot. Autom. **9**, 624–637 (1993)
20. Lederman, S.J., Abbott, S.G.: Texture perception: Studies of intersensory organization using a discrepancy paradigm, and visual versus tactile psychophysics. J. Exp. Psychol. Hum. Percept. Perform. **7**(4), 902–915 (1981)
21. Lederman, S., Thorne, G., Jones, B.: Perception of texture by vision and touch: Multidimensionality and intersensory integration. J. Exp. Psychol. Hum. Percept. Perform. **12**(2), 169–180 (1986)
22. Marks, L.E.: Cross-modal interactions in speeded classification. In: The Handbook of Multisensory Processes, pp. 85–105. MIT Press, Cambridge (2004)
23. Massaro, D.W.: Information-processing theory and strong inference: A paradigm for psychological inquiry. In: Perspectives on Perception and Action, pp. 273–299. Lawrence Erlbaum Associates, Hillsdale (1987)
24. Misceo, G.F., Hershberger, W.A., Mancini, R.L.: Haptic estimates of discordant visual-haptic size vary developmentally. Percept. Psychophys. **61**(4), 608–614 (1999)
25. Over, R.: An experimentally induced conflict between vision and propriozeption. Br. J. Psychol. **57**, 335–341 (1966)

26. Pick, H.L., Warren, D.H.: Sensory conflict in judgements of spatial direction. Percept. Psychophys. **6**(4), 203–205 (1969)
27. Rock, I., Harris, C.S.: Vision and touch. Sci. Am. **216**, 96–104 (1967)
28. Rock, I., Victor, J.: Vision and touch: An experimentally created conflict between the twosenses. Science **143**, 594–596 (1964)
29. Scheuchenpflug, R.: Measuring presence in virtual environments. In: Smith, M.J., Salvendey, G., Kasdorf, M.R. (eds.) HCI International 2001, New Orleans, pp. 56–58 (2001)
30. Sciavicco, L., Siciliano, B.: Modelling and Control of Robot Manipulators, 1st edn. Springer, New York (2003)
31. Shimojo, S., Shams, L.: Sensory modalities are not separate modalities: Plasticity and interactions. Curr. Opin. Neurobiol. **11**, 505–509 (2001)
32. Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., Kingstone, A.: The ventriloquist in motion: Illusory capture of dynamic information across sensory modalities. Cogn. Brain Res. **14**, 139–146 (2002)
33. Spence, C., Driver, J.: Crossmodal Space and Crossmodal Attention. Oxford University Press, Oxford (2004)
34. Spence, C., Nicholls, M.E.R., Driver, J.: The cost of expecting events in the wrong sensory modality. Percept. Psychophys. **63**(2), 330–336 (2001)
35. Srinivasan, M.A., Beauregard, G.L., Brock, D.L.: The impact of visual information on the haptic perception of stiffness in virtual environments. In: Proceedings of the ASME Dynamic Systems and Control Division—1996, DSC-vol. 58, pp. 555–559 (1996)
36. Stein, B.E., Meredith, M.A.: The Merging of the Senses. MIT Press, Cambridge (1993)
37. Überle, M., Buss, M.: Control of kinesthetic haptic interfaces. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (2004)
38. Warren, D.H., Pick, H.L.: Intermodality relations in localization in blind and sighted people. Percept. Psychophys. **8**(6), 430–432 (1970)
39. Welch, R.B., Warren, D.H.: Intersensory interactions. In: Handbook of Perception and Human Performance: Sensory Processes and Perception, vol. 1. Wiley, New York (1986)
40. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: A presence questionnaire. Presence: Teleoper. Virtual Environ. **7**(3), 225–240 (1998)
41. Wu, W.C., Basdogan, C., Srinivasan, M.A.: Visual, haptic, and bimodal perception of size and stiffness in virtual environments. In: Proceedings of the ASME Dynamic Systems and Control Division—1999, DSC-vol. 67, pp. 19–26 (1999)
42. Yokokohji, Y.: Bilateral control of master-slave manipulators for ideal kinesthetic coupling—formulation and experiment. IEEE Trans. Robot. Autom. **10**, 605–620 (1994)



**Martin Kuschel** is PhD candidate at the Institute of Automatic Control Engineering, Technische Universität (TU) München. He received his Dipl.-Ing. degree 2004 in electrical engineering from the Technische Universität Berlin. His research interests include cognition, human-machine interaction, visual–haptic perception, and robotics. He is currently spending a research fellowship at the Department of Psychology, Carnegie Mellon University in Pittsburgh, USA.

**Franziska Freyberger** received her doctorate in February 2008 at the Universität der Bundeswehr München. Her research interests include visual–haptic perception, the human role in human-machine interaction, and ergonomics. She is currently spending a sabbatical in Dublin, Ireland.



**Berthold Färber** is full professor, head of the Human Factors Institute, Faculty of Aerospace Engineering, Technische Universität der Bundeswehr München. His research interests include visual–haptic perception in telepresence systems and cognition to optimize driver assistance systems.

**Martin Buss** is full professor, head of the Institute of Automatic Control Engineering (LSR), Technische Universität München. Since 2006 he is the coordinator of the DFG Excellence Research Cluster *Cognition for Technical Systems—CoTeSys*. His research interests include automatic control, mechatronics, robotics, intelligent control, multimodal human–system interfaces, haptic (kinesthetic, tactile, thermal) systems, optimization, nonlinear and hybrid discrete-continuous systems, etc.