



**HAL**  
open science

# Visuo-Haptic Rendering of the Hand during 3D Manipulation in Augmented Reality

Erwan Normand, Claudio Pacchierotti, Eric Marchand, Maud Marchal

## ► To cite this version:

Erwan Normand, Claudio Pacchierotti, Eric Marchand, Maud Marchal. Visuo-Haptic Rendering of the Hand during 3D Manipulation in Augmented Reality. IEEE Transactions on Haptics (ToH), 2024, 17 (2), pp.277-291. <10.1109/TOH.2024.3358910>. <hal-04362252>

**HAL Id: hal-04362252**

**<https://hal.science/hal-04362252v1>**

Submitted on 22 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

# Visuo-Haptic Rendering of the Hand during 3D Manipulation in Augmented Reality

Erwan Normand, Claudio Pacchierotti, Eric Marchand, Maud Marchal

**Abstract**—Manipulating virtual objects with bare hands is a key interaction in Augmented Reality (AR) applications. However, there are still several limitations that affect the manipulation, including the lack of mutual visual occlusion between virtual and real content as well as the lack of haptic sensations. To address the two abovementioned matters, the role of the visuo-haptic rendering of the hand as sensory feedback is investigated. The first experiment explores the effect of showing the hand of the user as seen by the AR system through an avatar, comparing six visual hand rendering. The second experiment explores the effect of the visuo-haptic hand rendering by comparing two vibrotactile contact techniques provided at four delocalized positions on the hand and combined with the two most representative visual hand renderings from the first experiment. Results show that delocalized vibrotactile haptic hand rendering improved perceived effectiveness, realism, and usefulness when provided close to the contact point. However, the farthest rendering position, *i.e.*, on the contralateral hand, gave the best performance even though it was largely disliked. The visual hand rendering was perceived as less necessary for manipulation when the haptic hand rendering was available, but still provided useful feedback on the hand tracking.

**Index Terms**—Augmented Reality, Vibrotactile Feedback, Virtual Hands, Virtual Object Manipulation, Wearable Haptics.

## I. INTRODUCTION

**A**UGMENTED reality (AR) integrates virtual content into our real-world surroundings, giving the illusion of one unique environment and promising natural and seamless interactions with real and virtual objects. Virtual object manipulation is particularly critical for useful and effective AR usage, such as in medical applications, training, or entertainment [1], [2]. Hand tracking technologies [3], grasping techniques [4], and real-time physics engines permit users to directly manipulate virtual objects with their bare hands as if they were real [5], without requiring controllers [6], gloves [7], or predefined gesture techniques [8], [9]. Optical see-through AR (OST-AR) head-mounted displays (HMDs), such as the Microsoft HoloLens 2 or the Magic Leap, are particularly suited for this type of direct hand interaction [2].

However, there are still several haptic and visual limitations that affect manipulation in OST-AR, degrading the user experience. For example, it is difficult to estimate the position of one's hand in relation to a virtual content because mutual occlusion between the hand and the virtual object is often lacking [10], the depth of virtual content is underestimated [11], [12], and hand tracking still has a noticeable latency [3]. Similarly, it is challenging to ensure confident and

realistic contact with a virtual object due to the lack of haptic feedback and the intangibility of the virtual environment, which of course cannot apply physical constraints on the hand [13]–[16]. These limitations also make it difficult to confidently move a grasped object towards a target [13], [14].

To address these haptic and visual limitations, we investigate two types of sensory feedback that are known to improve virtual interactions with hands, but have not been studied together in an AR context: visual hand rendering and delocalized haptic rendering. A few works explored the effect of a visual hand rendering on interactions in AR by simulating mutual occlusion between the real hand and virtual objects [5], [9], [17], or displaying a 3D virtual hand model, semi-transparent [5], [9] or opaque [18]–[20]. Indeed, some visual hand renderings are known to improve interactions or user experience in virtual reality (VR), where the real hand is not visible [7], [21]–[24]. However, the role of a visual hand rendering superimposed and seen above the real tracked hand has not yet been investigated in AR. Conjointly, several studies have demonstrated that wearable haptics can significantly improve interactions performance and user experience in AR [13], [14], [25]. But haptic rendering for AR remains a challenge as it is difficult to provide rich and realistic haptic sensations while limiting their negative impact on hand tracking [26] and keeping the fingertips and palm free to interact with the real environment [15], [16], [25], [27]. Therefore, the haptic feedback of the fingertip contact with the virtual environment needs to be rendered elsewhere on the hand, it is unclear which positioning should be preferred or which type of haptic feedback is best suited for manipulating virtual objects in AR. A final question is whether one or the other of these (haptic or visual) hand renderings should be preferred [13], [14], or whether a combined visuo-haptic rendering is beneficial for users. In fact, both hand renderings can provide sufficient sensory cues for efficient manipulation of virtual objects in AR, or conversely, they can be shown to be complementary.

In this paper, we investigate the role of the visuo-haptic rendering of the hand during 3D manipulation of virtual objects in OST-AR. We consider two representative manipulation tasks: push-and-slide and grasp-and-place a virtual object. The main contributions of this work are:

- a first human subject experiment evaluating the performance and user experience of six visual hand renderings superimposed on the real hand;
- a second human subject experiment evaluating the performance and user experience of visuo-haptic hand renderings by comparing two vibrotactile contact techniques provided at four delocalized positions on the hand and combined with the two most representative visual hand renderings established in the first experiment.

E. Normand and E. Marchand are with Univ Rennes, Inria, CNRS, IRISA (Rennes, France). E-mail: {e.normand, e.marchand}@irisa.fr

C. Pacchierotti is with CNRS, Univ Rennes, Inria, IRISA (Rennes, France). E-mail: claudio.pacchierotti@irisa.fr

M. Marchal is with Univ Rennes, INSA Rennes, IRISA, CNRS, Inria (Rennes, France) and IUF. E-mail: maud.marchal@irisa.fr

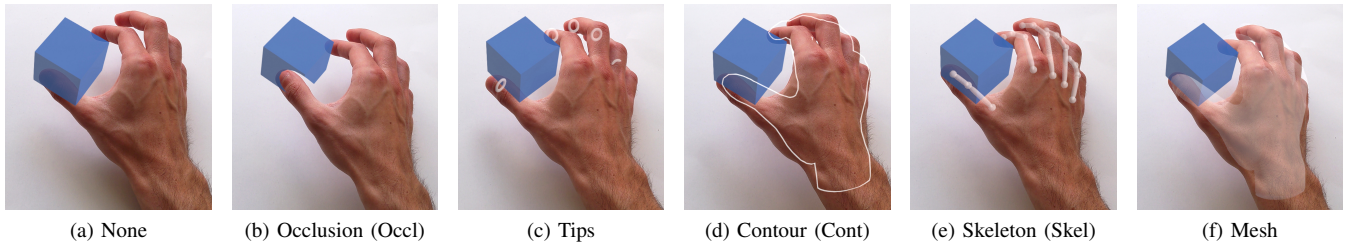


Figure 1. Experiment #1. The six considered visual hand renderings, as seen by the user through the AR headset during the two-finger grasping of a virtual cube. From left to right: no visual rendering (*None*), cropped virtual content to enable hand-cube occlusion (*Occlusion, Occl*), rings on the fingertips (*Tips*), thin outline of the hand (*Contour, Cont*), fingers’ joints and phalanges (*Skeleton, Skel*), and semi-transparent 3D hand model (*Mesh*).

## II. RELATED WORK

This Section summarizes the state of the art in visual hand rendering and (wearable) haptic rendering in AR, focusing on virtual object manipulation.

### A. Visual Hand Rendering in AR

Mutual visual occlusion between a virtual object and the real hand, *i.e.*, hiding the virtual object when the real hand is in front of it and hiding the real hand when it is behind the virtual object, is often presented as natural and realistic, enhancing the blending of real and virtual environments [5], [17]. In video see-through AR (VST-AR), this could be solved as a masking problem by combining the image of the real world captured by a camera and the generated virtual image [10]. In OST-AR, this is more difficult because the virtual environment is displayed as a transparent 2D image on top of the 3D real world, which cannot be easily masked [10]. Moreover, in VST-AR, the grip aperture and depth positioning of virtual objects often seem to be wrongly estimated [13], [17]. However, this effect has yet to be verified in an OST-AR setup.

An alternative is to render the virtual objects and the hand semi-transparently, so that they are partially visible even when one is occluding the other, *e.g.*, in Fig. 1a the real hand is behind the virtual cube but still visible. Although perceived as less natural, this seems to be preferred to a mutual visual occlusion in VST-AR [5], [9], [28] and VR [24], but has not yet been evaluated in OST-AR. However, this effect still causes depth conflicts that make it difficult to determine if one’s hand is behind or in front of a virtual object, *e.g.*, in Fig. 1a the thumb is in front of the virtual cube, but it appears to be behind it.

In VR, as the user is fully immersed in the virtual environment and cannot see their real hands, it is necessary to represent them virtually. It is known that the virtual hand representation has an impact on perception, interaction performance, and preference of users [7], [21]–[23]. In a pick-and-place task in VR, Prachyabrued *et al.* [7] found that the virtual hand representation whose motion was constrained to the surface of the virtual objects performed the worst, while the virtual hand representation following the tracked human hand (thus penetrating the virtual objects), performed the best, even though it was rather disliked. The authors also observed that the best compromise was a double rendering, showing both the tracked hand and a hand rendering constrained by the virtual

environment. It has also been shown that over a realistic avatar, a skeleton rendering (similar to Fig. 1e) can provide a stronger sense of being in control [21] and that minimalistic fingertip rendering (similar to Fig. 1c) can be more effective in a typing task [22].

In AR, as the real hand of a user is visible but not physically constrained by the virtual environment, adding a visual hand rendering that can physically interact with virtual objects would achieve a similar result to the promising double-hand rendering of Prachyabrued *et al.* [7]. Additionally, Kahl *et al.* [29] showed that a virtual object overlaying a tangible object in OST-AR can vary in size without worsening the users’ experience nor the performance. This suggests that a visual hand rendering superimposed on the real hand could be helpful, but should not impair users.

Few works have explored the effect of visual hand rendering in AR [6], [13], [18]–[20]. For example, Blaga *et al.* [18] evaluated a skeleton rendering in several virtual object manipulations against no visual hand overlay. Performance did not improve, but participants felt more confident with the virtual hand. However, the experiment was carried out on a screen, in a non-immersive AR scenario. Saito *et al.* [20] found that masking the real hand with a textured 3D opaque virtual hand did not improve performance in a reach-to-grasp task but displaying the points of contact on the virtual object did. To the best of our knowledge, evaluating the role of a visual rendering of the hand displayed “and seen” directly above real tracked hands in immersive OST-AR has not been explored, particularly in the context of virtual object manipulation.

### B. Wearable Haptic Feedback in AR

Different haptic feedback systems have been explored to improve interactions in AR, including grounded force feedback devices [30]–[32], exoskeletons [33], tangible objects [3], [34]–[37], and wearable haptic devices [15], [16], [26], [38].

Wearable haptics seems particularly suited for this context, as it takes into account many of the AR constraints, *e.g.*, limited impact on hand tracking performance and reduced impairment of the senses and ability of the users to interact with real content [13]–[16], [26], [38]–[40]. For example, Pacchierotti *et al.* [26] designed a haptic ring providing pressure and skin stretch sensations to be worn at the proximal finger phalanx, so as to improve the hand tracking during a pick-and-place task. Pezent *et al.* [38] proposed Tasbi: a wristband haptic device capable of rendering vibrations and

pressures. Teng *et al.* [16] presented Touch&Fold, a haptic device attached to the nail that provides pressure and texture sensations when interacting with virtual content, but also folds away when the user interacts with real objects, leaving the fingertip free. This approach was also perceived as more realistic than providing sensations directly on the nail, as in [41]. Each of these haptic devices provided haptic feedback about fingertip interactions with the virtual content on other parts of the hand. If it is indeed necessary to delocalize the haptic feedback, each of these positions is promising, and they have not yet been compared with each other.

Conjointly, a few studies have explored and compared the effects of visual and haptic feedback in tasks involving the manipulation of virtual objects with the hand. Sarac *et al.* [25] and Palmer *et al.* [27] studied the effects of providing haptic feedback about contacts at the fingertips using haptic devices worn at the wrist, testing different mappings. Results proved that moving the haptic feedback away from the point(s) of contact is possible and effective, and that its impact is more significant when the visual feedback is limited. In pick-and-place tasks in AR involving both virtual and real objects, Maisto *et al.* [13] and Meli *et al.* [14] showed that having a haptic rendering of the fingertip interactions with the virtual objects led to better performance and perceived effectiveness than having only a visual rendering of the hand, similar to Fig. 1c. Moreover, employing the haptic ring of [26] on the proximal finger phalanx led to an improved performance with respect to more standard fingertip haptic devices [42]. However, the measured difference in performance could be attributed to either the device or the device position (proximal vs fingertip), or both. Furthermore, all of these studies were conducted in non-immersive setups, where users looked at a screen displaying the visual interactions, and only compared haptic and visual feedback, but did not examine them together. The improved performance and perceived effectiveness of a delocalized haptic feedback over a visual feedback alone, or their multimodal combination, remains to be verified in an immersive OST-AR setup.

These limitations have motivated the research presented in this paper, where we aim at filling the abovementioned gaps related to the understanding of the role of the visuo-haptic rendering of the hand during the 3D manipulation of virtual objects in AR.

### III. EXPERIMENT #1: VISUAL RENDERING OF THE HAND IN AR

This first experiment aims to analyze whether the chosen visual hand rendering affects the performance and user experience of manipulating virtual objects with bare hands in AR.

#### A. Visual Hand Renderings

We compared a set of the most popular visual hand renderings, as also presented in Sec. II-A. Since we address hand-centered manipulation tasks, we only considered renderings including the fingertips. Moreover, as to keep the focus on the

hand rendering itself, we used neutral semi-transparent grey meshes, consistent with the choices made in [19], [24]. All considered hand renderings are drawn following the tracked pose of the user's real hand. However, while the real hand can of course penetrate virtual objects, the visual hand is always constrained by the virtual environment.

1) *None (Fig. 1a)*: As a reference, we considered no visual hand rendering, as is common in AR [3], [16], [18], [34]. Users have no information about hand tracking and no feedback about contact with the virtual objects, other than their movement when touched. As virtual content is rendered on top of the real environment, the hand of the user can be hidden by the virtual objects when manipulating them (see Sec. II-A).

2) *Occlusion (Occl, Fig. 1b)*: To avoid the abovementioned undesired occlusions due to the virtual content being rendered on top of the real environment, we can carefully crop the former whenever it hides real content that should be visible [10], e.g., the thumb of the user in Fig. 1b. This approach is frequent in works using VST-AR headsets [5], [9], [17], [32], [43].

3) *Tips (Fig. 1c)*: This rendering shows small visual rings around the fingertips of the user, highlighting the most important parts of the hand and contact with virtual objects during fine manipulation. Unlike work using small spheres [13], [22], [23], [37], [44], this ring rendering also provides information about the orientation of the fingertips.

4) *Contour (Cont, Fig. 1d)*: This rendering is a 1-mm-thick outline contouring the user's hands, providing information about the whole hand while leaving its inside visible. Unlike the other renderings, it is not occluded by the virtual objects, as shown in Fig. 1d. This rendering is not as usual as the previous others in the literature [45].

5) *Skeleton (Skel, Fig. 1e)*: This rendering schematically renders the joints and phalanges of the fingers with small spheres and cylinders, respectively, leaving the outside of the hand visible. It can be seen as an extension of the Tips rendering to include the complete fingers articulations. It is widely used in VR [21], [23], [46] and AR [18], [19], as it is considered simple yet rich and comprehensive.

6) *Mesh (Fig. 1f)*: This rendering is a 3D semi-transparent ( $\alpha = 0.2$ ) hand model, which is common in VR [7], [19], [21], [23], [24], [46]. It can be seen as a filled version of the Contour hand rendering, thus partially covering the view of the real hand.

#### B. Manipulation Tasks and Virtual Scene

Following the guidelines of Bergström *et al.* [47] for designing object manipulation tasks, we considered two variations of a 3D pick-and-place task, commonly found in interaction and manipulation studies [7], [13], [14], [18], [24].

1) *Push Task*: The first manipulation task consists in pushing a virtual object along a real flat surface towards a target placed on the same plane (see Fig. 2a). The virtual object to manipulate is a small 50 mm blue and opaque cube, while the target is a (slightly) bigger 70 mm blue and semi-transparent volume. At every repetition of the task, the cube to manipulate always spawns at the same place, on top of a real table in front

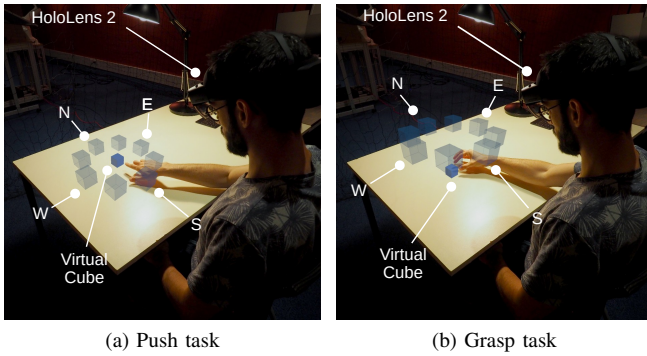


Figure 2. Experiment #1. The two manipulation tasks: (a) pushing a virtual cube along a table towards a target placed on the same surface; (b) grasping and lifting a virtual cube towards a target placed on a 20-cm-higher plane. Both pictures show the cube to manipulate in the middle (5-cm-edge and opaque) and the eight possible targets to reach (7-cm-edge volume and semi-transparent). Only one target at a time was shown during the experiments.

of the user. On the other hand, the target volume can spawn in eight different locations on the same table, located on a 20 cm radius circle centered on the cube, at  $45^\circ$  from each other (see again Fig. 2a). Users are asked to push the cube towards the target volume using their fingertips in any way they prefer. In this task, the cube cannot be lifted. The task is considered completed when the cube is *fully* inside the target volume.

2) *Grasp Task*: The second manipulation task consists in grasping, lifting, and placing a virtual object in a target placed on a different (higher) plane (see Fig. 2b). The cube to manipulate and target volume are the same as in the previous task. However, this time, the target volume can spawn in eight different locations on a plane 10 cm *above* the table, still located on a 20 cm radius circle at  $45^\circ$  from each other. Users are asked to grasp, lift, and move the cube towards the target volume using their fingertips in any way they prefer. As before, the task is considered completed when the cube is *fully* inside the volume.

### C. Experimental Design

We analyzed the two tasks separately. For each of them, we considered two independent, within-subject, variables:

- *Visual Hand Renderings*, consisting of the six possible renderings discussed in Sec. III-A: None, Occlusion (Occl), Tips, Contour (Cont), Skeleton (Skel), and Mesh.
- *Target*, consisting of the eight possible location of the target volume, named as the cardinal points and as shown in Fig. 2: E, NE, N, NW, W, SW, S, and SE.

Each condition was repeated three times. To control learning effects, we counter-balanced the orders of the two manipulation tasks and visual hand renderings following a  $6 \times 6$  Latin square, leading to six blocks where the position of the target volume was in turn randomized. This design led to a total of 2 manipulation tasks  $\times$  6 visual hand renderings  $\times$  8 targets  $\times$  3 repetitions = 288 trials per participant.

### D. Apparatus and Implementation

We used the OST-AR headset HoloLens 2. It is capable of rendering virtual content within an horizontal field of view

of  $43^\circ$  and a vertical one of  $29^\circ$ . It is also able to track the environment as well as the user's fingers. We measured the latency of the hand tracking at 15 ms, independent of the hand movement speed.

The implementation of our experiment was done in C# using Unity 2022.1, PhysX 4.1, and the Mixed Reality Toolkit (MRTK) 2.8<sup>1</sup>. The compiled application ran directly on the HoloLens 2 at 60 FPS.

The default 3D hand model from MRTK was used for all visual hand renderings. By changing the material properties of this hand model, we were able to achieve the six renderings shown in Fig. 1. A calibration was performed for every participant, so as to best adapt the size of the visual hand rendering to their real hand. A set of empirical tests enabled us to choose the best rendering characteristics in terms of transparency and brightness for the virtual objects and hand renderings, which were applied throughout the experiment.

The hand tracking information provided by MRTK was used to construct a virtual articulated physics-enabled hand using PhysX. It featured 25 DoFs, including the fingers proximal, middle, and distal phalanges. To allow effective (and stable) physical interactions between the hand and the virtual cube to manipulate, we implemented an approach similar to that of Borst *et al.* [48], where a series of virtual springs with high stiffness are used to couple the physics-enabled hand with the tracked hand. As before, a set of empirical tests have been used to select the most effective physical characteristics in terms of mass, elastic constant, friction, damping, colliders size, and shape for the (tracked) virtual hand interaction model.

The room where the experiment was held had no windows, with one light source of 800 lm placed 70 cm above the table. This setup enabled a good and consistent tracking of the user's fingers.

### E. Protocol

First, participants were given a consent form that briefed them about the tasks and the protocol of the experiment. Then, participants were asked to comfortably sit in front of a table and wear the HoloLens 2 headset as shown in Fig. 2, perform the calibration of the visual hand size as described in Sec. III-D, and complete a 2-minutes training to familiarize with the AR rendering and the two considered tasks. During this training, we did not use any of the six hand renderings we want to test, but rather a fully-opaque white hand rendering that completely occluded the real hand of the user.

Participants were asked to carry out the two tasks as naturally and as fast as possible. Similarly to [7], [13], [18], [24], we only allowed the use of the dominant hand. The experiment took around 1 hour and 20 minutes to complete.

### F. Participants

Twenty-four subjects participated in the study (eight aged between 18 and 24, fourteen aged between 25 and 34, and two aged between 35 and 44; 22 males, 1 female, 1 preferred not to say). None of the participants reported any deficiencies

<sup>1</sup> <https://learn.microsoft.com/windows/mixed-reality/mrtk-unity>

in their visual perception abilities. Two subjects were left-handed, while the twenty-two other were right-handed; they all used their dominant hand during the trials. Ten subjects had significant experience with VR (“I use it every week”), while the fourteen other reported little to no experience with VR. Two subjects had significant experience with AR (“I use it every week”), while the twenty-two other reported little to no experience with AR. Participants signed an informed consent, including the declaration of having no conflict of interest.

### G. Collected Data

Inspired by LaViola Jr *et al.* [1], we collected the following metrics during the experiment. (i) The task *Completion Time*, defined as the time elapsed between the very first contact with the virtual cube and its correct placement inside the target volume; as subjects were asked to complete the tasks as fast as possible, lower completion times mean better performance. (ii) The number of *Contacts*, defined as the number of separate times the user’s hand makes contact with the virtual cube; in both tasks, a lower number of contacts means a smoother continuous interaction with the object. Finally, (iii) the mean *Time per Contact*, defined as the total time any part of the user’s hand contacted the cube divided by the number of contacts; higher values mean that the user interacted with the object for longer non-interrupted periods of time. Solely for the grasp-and-place task, we also measured the (iv) *Grip Aperture*, defined as the average distance between the thumb’s fingertip and the other fingertips during the grasping of the cube; lower values indicate a greater finger interpenetration with the cube, resulting in a greater discrepancy between the real hand and the visual hand rendering constrained to the cube surfaces and showing how confident users are in their grasp [7], [17], [18], [46]. Taken together, these measures provide an overview of the performance and usability of each of the visual hand renderings tested, as we hypothesized that they should influence the behavior and effectiveness of the participants.

At the end of each task, participants were asked to rank the visual hand renderings according to their preference with respect to the considered task. Participants also rated each visual hand rendering individually on six questions using a 7-item Likert scale (1=Not at all, 7=Extremely): (*Difficulty*) How difficult were the tasks? (*Fatigue*) How fatiguing (mentally and physically) were the tasks? (*Precision*) How precise were you in performing the tasks? (*Efficiency*) How fast/efficient do you think you were in performing the tasks? (*Rating*) How much do you like each visual hand?

Finally, participants were encouraged to comment out loud on the conditions throughout the experiment, as well as in an open-ended question at its end, so as to gather additional qualitative information.

### H. Results

Results of each trials measure were analyzed with a linear mixed model (LMM), with the order of the two manipulation tasks and the six visual hand renderings (Order), the visual hand renderings (Hand), the target volume position (Target),

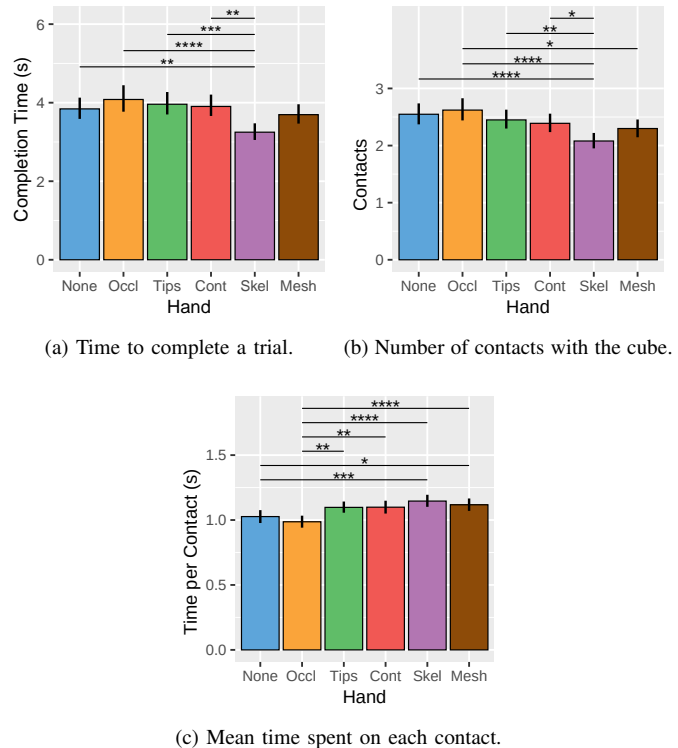


Figure 3. Experiment #1: Push task. Geometric means with bootstrap 95 % confidence interval for each visual hand rendering and Tukey’s HSD pairwise comparisons: \*\*\* is  $p < 0.001$ , \*\* is  $p < 0.01$ , and \* is  $p < 0.05$ .

and their interactions as fixed effects and the Participant as random intercept. For every LMM, residuals were tested with a Q-Q plot to confirm normality. On statistically significant effects, estimated marginal means of the LMM were compared pairwise using Tukey’s HSD test. Only significant results were reported.

Because Completion Time, Contacts, and Time per Contact measure results were Gamma distributed, they were first transformed with a log to approximate a normal distribution. Their analysis results are reported anti-logged, corresponding to geometric means of the measures.

#### 1) Push Task:

**Completion Time:** On the time to complete a trial, there were two statistically significant effects: Hand ( $F_{5,2868} = 24.8$ ,  $p < 0.001$ , see Fig. 3b) and Target ( $F_{7,2868} = 5.9$ ,  $p < 0.001$ ). Skeleton was the fastest, more than None (+18 %,  $p = 0.005$ ), Occlusion (+26 %,  $p < 0.001$ ), Tips (+22 %,  $p < 0.001$ ), and Contour (+20 %,  $p = 0.001$ ). Three groups of targets volumes were identified: (1) sides E, W, and SW targets were the fastest; (2) back and front NE, S, and SE were slower ( $p = 0.003$ ); and (3) back N and NW targets were the slowest ( $p = 0.04$ ).

**Contacts:** On the number of contacts, there were two statistically significant effects: Hand ( $F_{5,2868} = 6.7$ ,  $p < 0.001$ , see Fig. 3b) and Target ( $F_{7,2868} = 27.8$ ,  $p < 0.001$ ). Fig. 3b shows the Contacts for each Hand. Less contacts were made with Skeleton than with None (−23 %,  $p < 0.001$ ), Occlusion (−26 %,  $p < 0.001$ ), Tips (−18 %,  $p = 0.004$ ), and Contour (−15 %,  $p = 0.02$ ); and less with Mesh than with Occlusion

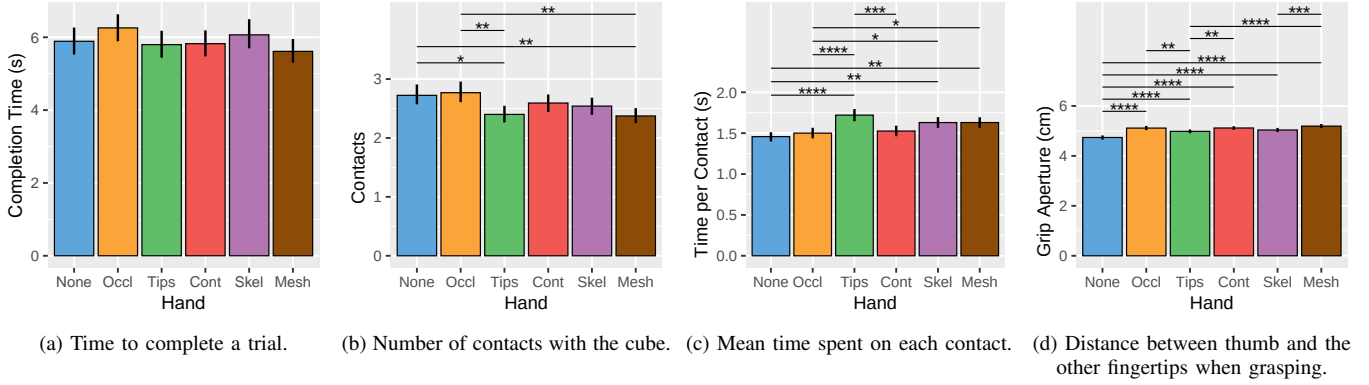


Figure 4. Experiment #1: Grasp task. Geometric means with bootstrap 95 % confidence interval for each visual hand rendering and Tukey's HSD pairwise comparisons: \*\*\* is  $p < 0.001$ , \*\* is  $p < 0.01$ , and \* is  $p < 0.05$ .

(-14 %,  $p = 0.04$ ). This indicates how effective a visual hand rendering is: a lower result indicates a smoother ability to push and rotate properly the cube into the target, as one would probably do with a real cube. Targets on the left (W) and the right (E, SW) were easier to reach than the back ones (N, NW,  $p < 0.001$ ).

**Time per Contact:** On the mean time spent on each contact, there were two statistically significant effects: Hand ( $F_{5,2868} = 8.4$ ,  $p < 0.001$ , see Fig. 3c) and Target ( $F_{7,2868} = 19.4$ ,  $p < 0.001$ ). It was shorter with None than with Skeleton (-10 %,  $p < 0.001$ ) and Mesh (-8 %,  $p = 0.03$ ); and shorter with Occlusion than with Tips (-10 %,  $p = 0.002$ ), Contour (-10 %,  $p = 0.001$ ), Skeleton (-14 %, 0.001), and Mesh (-12 %,  $p = 0.03$ ). This result suggests that users pushed the virtual cube with more confidence with a visible visual hand rendering. On the contrary, the lack of visual hand constrained the participants to give more attention to the cube's reactions. Targets on the left (W, SW) and the right (E) sides had higher Timer per Contact than all the other targets ( $p = 0.005$ ).

## 2) Grasp Task:

**Completion Time:** On the time to complete a trial, there was one statistically significant effect of Target ( $F_{7,2868} = 37.2$ ,  $p < 0.001$ ) but not of Hand ( $F_{5,2868} = 1.8$ ,  $p = 0.1$ , see Fig. 4a). Targets on the back and the left (N, NW, and W) were slower than targets on the front (SW, S, and SE,  $p = 0.003$ ) except for NE (back-right) which was also fast.

**Contacts:** On the number of contacts, there were two statistically significant effects: Hand ( $F_{5,2868} = 5.2$ ,  $p < 0.001$ , see Fig. 4b) and Target ( $F_{7,2868} = 21.2$ ,  $p < 0.001$ ). Less contacts were made with Tips than with None (-13 %,  $p = 0.02$ ) and Occlusion (-15 %,  $p = 0.004$ ); and less with Mesh than with None (-15 %,  $p = 0.006$ ) and Occlusion (-17 %,  $p = 0.001$ ). This result suggests that having no visible visual hand increased the number of failed grasps or cube drops. But, surprisingly, only Tips and Mesh were statistically significantly better, not Contour nor Skeleton. Targets on the back and left were more difficult (N, NW, and W) than targets on the front (SW, S, and SE,  $p < 0.001$ ).

**Time per Contact:** On the mean time spent on each contact, there were two statistically significant effects: Hand ( $F_{5,2868} = 9.6$ ,  $p < 0.001$ , see Fig. 4c) and Target ( $F_{7,2868} =$

5.6,  $p < 0.001$ ). It was shorter with None than with Tips (-15 %,  $p < 0.001$ ), Skeleton (-11 %,  $p = 0.001$ ) and Mesh (-11 %,  $p = 0.001$ ); shorter with Occlusion than with Tips (-10 %,  $p < 0.001$ ), Skeleton (-8 %,  $p = 0.05$ ), and Mesh (-8 %,  $p = 0.04$ ); shorter with Contour than with Tips (-8 %,  $p < 0.001$ ). As for the Push task, the lack of visual hand increased the number of failed grasps or cube drops. The Tips rendering seemed to provide one of the best feedback for the grasping, maybe thanks to the fact that it provides information about both position and rotation of the tracked fingertips. This time was the shortest on the front S than on the other target volumes ( $p < 0.001$ ).

**Grip Aperture:** On the average distance between the thumb's fingertip and the other fingertips during grasping, there were two statistically significant effects: Hand ( $F_{5,2868} = 35.8$ ,  $p < 0.001$ , see Fig. 4d) and Target ( $F_{7,2868} = 3.7$ ,  $p < 0.001$ ). It was shorter with None than with Occlusion ( $p < 0.001$ ), Tips ( $p < 0.001$ ), Contour ( $p < 0.001$ ), Skeleton ( $p < 0.001$ ) and Mesh ( $p < 0.001$ ); shorter with Tips than with Occlusion ( $p = 0.008$ ), Contour ( $p = 0.006$ ) and Mesh ( $p < 0.001$ ); and shorter with Skeleton than with Mesh ( $p < 0.001$ ). This result is an evidence of the lack of confidence of participants with no visual hand rendering: they grasped the cube more to secure it. The Mesh rendering seemed to have provided the most confidence to participants, maybe because it was the closest to the real hand. The Grip Aperture was longer on SE (bottom-right) target volume, indicating a higher confidence, than on back and side targets (E, NE, N, W,  $p = 0.03$ ).

**3) Ranking:** Fig. 5 shows the ranking of each visual hand rendering for the Push and Grasp tasks. Friedman tests indicated that both ranking had statistically significant differences ( $p < 0.001$ ). Pairwise Wilcoxon signed-rank tests with Holm-Bonferroni adjustment were then used on both ranking results (see Sec. III-G):

- **Push Ranking:** Occlusion was ranked lower than Contour ( $p = 0.005$ ), Skeleton ( $p = 0.02$ ), and Mesh ( $p = 0.03$ ); Tips was ranked lower than Skeleton ( $p = 0.02$ ). This good ranking of the Skeleton rendering for the Push task is consistent with the Push trial results.
- **Grasp Ranking:** Occlusion was ranked lower than Con-

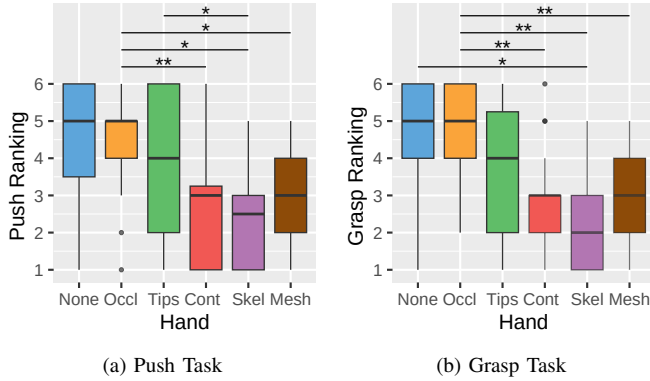


Figure 5. Experiment #1. Boxplots of the ranking (lower is better) of each visual hand rendering and pairwise Wilcoxon signed-rank tests with Holm-Bonferroni adjustment: \*\* is  $p < 0.01$  and \* is  $p < 0.05$ .

tour ( $p = 0.001$ ), Skeleton ( $p = 0.001$ ), and Mesh ( $p = 0.007$ ); No Hand was ranked lower than Skeleton ( $p = 0.04$ ). A complete visual hand rendering seemed to be preferred over no visual hand rendering when grasping.

4) *Questionnaire*: Fig. 6 presents the questionnaire results for each visual hand rendering. Friedman tests indicated that all questions had statistically significant differences ( $p < 0.001$ ). Pairwise Wilcoxon signed-rank tests with Holm-Bonferroni adjustment were then used each question results (see Sec. III-G):

- *Difficulty*: Occlusion was considered more difficult than Contour ( $p = 0.02$ ), Skeleton ( $p = 0.01$ ), and Mesh ( $p = 0.03$ ).
- *Fatigue*: None was found more fatiguing than Mesh ( $p = 0.04$ ); And Occlusion more than Skeleton ( $p = 0.02$ ) and Mesh ( $p = 0.02$ ).
- *Precision*: None was considered less precise than Skeleton ( $p = 0.02$ ) and Mesh ( $p = 0.02$ ); And Occlusion more than Contour ( $p = 0.02$ ), Skeleton ( $p = 0.006$ ), and Mesh ( $p = 0.02$ ).
- *Efficiency*: Occlusion was found less efficient than Contour ( $p = 0.01$ ), Skeleton ( $p = 0.02$ ), and Mesh ( $p = 0.02$ ).
- *Rating*: Occlusion was rated lower than Contour ( $p = 0.02$ ) and Skeleton ( $p = 0.03$ ).

In summary, Occlusion was worse than Skeleton for all questions, and worse than Contour and Mesh on 5 over 6 questions. Results of Difficulty, Performance, and Precision questions are consistent in that way. Moreover, having no visible visual hand rendering was felt by users fatiguing and less precise than having one. Surprisingly, no clear consensus was found on Rating. Each visual hand rendering, except for Occlusion, had simultaneously received the minimum and maximum possible notes.

## I. Discussion

We evaluated six visual hand renderings, as described in Sec. III-A, displayed on top of the real hand, in two virtual object manipulation tasks in AR.

During the Push task, the Skeleton hand rendering was the fastest (see Fig. 3a), as participants employed fewer and longer contacts to adjust the cube inside the target volume (see Fig. 3b and Fig. 3c). Participants consistently used few and continuous contacts for all visual hand renderings (see Fig. 3b), with only less than ten trials, carried out by two participants, quickly completed with multiple discrete touches. However, during the Grasp task, despite no difference in completion time, providing no visible hand rendering (None and Occlusion renderings) led to more failed grasps or cube drops (see Fig. 4a and Fig. 4c). Indeed, participants found the None and Occlusion renderings less effective (see Fig. 5b) and less precise (see Fig. 6). To understand whether the participants' previous experience might have played a role, we also carried out an additional statistical analysis considering VR experience as an additional between-subjects factor, *i.e.*, VR novices vs. VR experts ("I use it every week", see Sec. III-F). We found no statistically significant differences when comparing the considered metrics between VR novices and experts.

Interestingly, all visual hand renderings showed grip apertures very close to the size of the virtual cube, except for the None rendering (see Fig. 4d), with which participants applied stronger grasps, *i.e.*, less distance between the fingertips. Having no visual hand rendering, but only the reaction of the cube to the interaction as feedback, made participants less confident in their grip. This result contrasts with the wrongly estimated grip apertures observed by Al-Kalbani *et al.* [17] in an exocentric VST-AR setup. Also, while some participants found the absence of visual hand rendering more natural, many of them commented on the importance of having feedback on the tracking of their hands, as observed by Xiao *et al.* [3] in a similar immersive OST-AR setup.

Yet, participants' opinions of the visual hand renderings were mixed on many questions, except for the Occlusion one, which was perceived less effective than more "complete" visual hands such as Contour, Skeleton, and Mesh hands (see Fig. 6). However, due to the latency of the hand tracking and the visual hand reacting to the cube, almost all participants thought that the Occlusion rendering to be a "shadow" of the real hand on the cube.

The Tips rendering, which showed the contacts made on the virtual cube, was controversial as it received the minimum and the maximum score on every question. Many participants reported difficulties in seeing the orientation of the visual fingers, while others found that it gave them a better sense of the contact points and improved their concentration on the task. This result are consistent with Saito *et al.* [20], who found that displaying the points of contacts was beneficial for grasping a virtual object over an opaque visual hand overlay.

To summarize, when employing a visual hand rendering overlaying the real hand, participants were more performant and confident in manipulating virtual objects with bare hands in AR. These results contrast with similar manipulation studies, but in non-immersive, on-screen AR, where the presence of a visual hand rendering was found by participants to improve the usability of the interaction, but not their performance [13], [14], [18]. Our results show the most effective visual hand rendering to be the Skeleton one. Participants

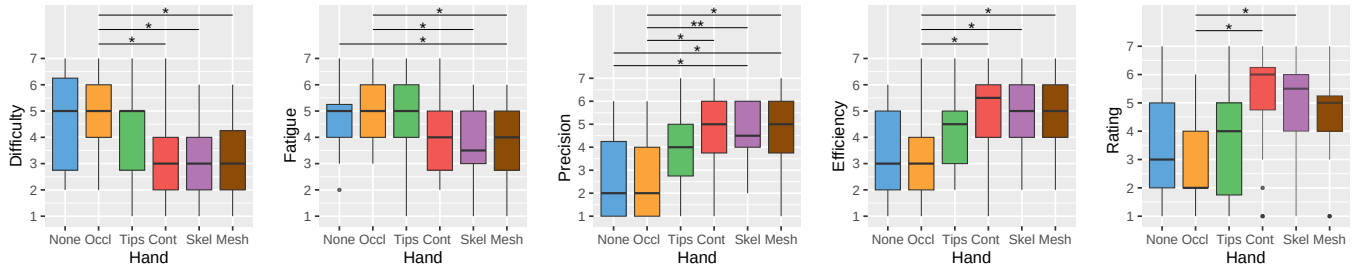


Figure 6. Experiment #1. Boxplots of the questionnaire results of each visual hand rendering and pairwise Wilcoxon signed-rank tests with Holm-Bonferroni adjustment: \*\* is  $p < 0.01$  and \* is  $p < 0.05$ . Lower is better for Difficulty and Fatigue. Higher is better for Precision, Efficiency, and Rating.

appreciated that it provided a detailed and precise view of the tracking of the real hand, without hiding or masking it. Although the Contour and Mesh hand renderings were also highly rated, some participants felt that they were too visible and masked the real hand. This result is in line with the results of virtual object manipulation in VR of Prachyabrued *et al.* [7], who found that the most effective visual hand rendering was a double representation of both the real tracked hand and a visual hand physically constrained by the virtual environment. This type of Skeleton rendering was also the one that provided the best sense of agency (control) in VR [21], [23].

These results have of course some limitations as they only address limited types of manipulation tasks and visual hand characteristics, evaluated in a specific OST-AR setup. The two manipulation tasks were also limited to placing a virtual cube in predefined target volumes. Testing a wider range of virtual objects and more ecological tasks *e.g.*, stacking, assembly, will ensure a greater applicability of the results obtained in this work, as well as considering bimanual manipulation. Similarly, a broader experimental study might shed light on the role of gender and age, as our subject pool was not sufficiently diverse in this respect. However, we believe that the results presented here provide a rather interesting overview of the most promising approaches in AR manipulation.

#### IV. EXPERIMENT #2: VISUO-HAPTIC RENDERING OF THE HAND IN AR

Providing haptic feedback during free-hand manipulation in AR is not trivial, as wearing haptic devices on the hand might affect the tracking capabilities of the system. Moreover, it is important to leave the user capable of interacting with both virtual and real objects, avoiding the use of haptic interfaces that cover the fingertips or palm. For this reason, it is often considered beneficial to move the point of application of the haptic rendering elsewhere on the hand (see Sec. II-B).

This second experiment aims to evaluate whether a visuo-haptic hand rendering affects the performance and user experience of manipulation of virtual objects with bare hands in AR. The chosen visuo-haptic hand renderings are the combination of the two most representative visual hand renderings established in the first experiment, *i.e.*, Skeleton and None, described in Sec. III-A, with two contact vibration techniques provided at four delocalized positions on the hand.

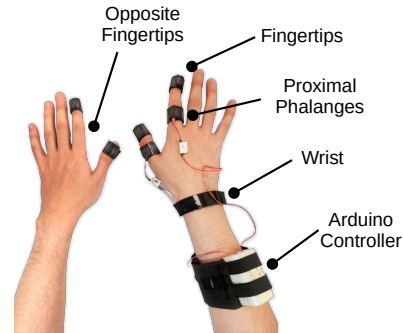


Figure 7. Experiment #2: setup of the vibrotactile devices. To ensure minimal encumbrance, we used the same two motors throughout the experiment, moving them to the considered positioning before each new experimental block (in this case, on the co-located proximal phalanges, *Prox*). Thin self-gripping straps were placed on the five considered positionings during the entirety of the experiment.

##### A. Vibrotactile Renderings

The vibrotactile hand rendering provided information about the contacts between the virtual object and the thumb and index fingers of the user, as they were the two fingers most used for grasping in our first experiment. We evaluated both the delocalized positioning and the contact vibration technique of the vibrotactile hand rendering.

###### 1) Vibrotactile Positionings:

- *Fingertips (Tips)*: Vibrating actuators were placed right above the nails, similarly to [41]. This is the positioning closest to the fingertips.
- *Proximal Phalanges (Prox)*: Vibrating actuators were placed on the dorsal side of the proximal phalanges, similarly to [13], [14], [42].
- *Wrist (Wris)*: Vibrating actuators providing contacts rendering for the index and thumb were placed on ulnar and radial sides of the wrist, similarly to [25], [27], [38].
- *Opposite fingertips (Oppo)*: Vibrating actuators were placed on the fingertips of contralateral hand, also above the nails, similarly to [35], [49].
- *Nowhere (Nowh)*: As a reference, we also considered the case where we provided no vibrotactile rendering.

2) *Contact Vibration Techniques*: When a fingertip contacts the virtual cube, we activate the corresponding vibrating actuator. We considered two representative contact vibration

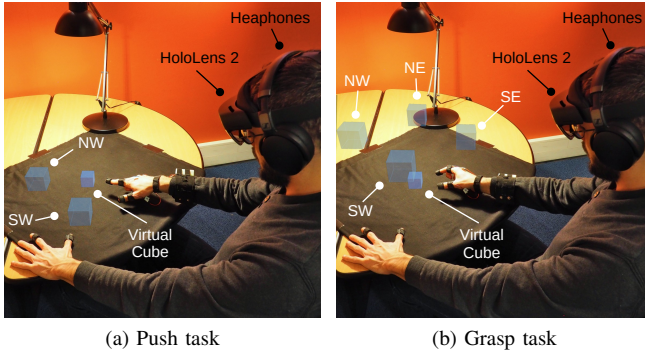


Figure 8. Experiment #2. The two manipulation tasks: (a) pushing a virtual cube along a table toward a target placed on the same surface; (b) grasping and lifting a virtual cube toward a target placed on a 20 cm higher plane. Both pictures show the cube to manipulate in the middle (5 cm and opaque) and the eight possible targets to reach (7 cm cube and semi-transparent). Only one target at a time was shown during the experiments.

techniques, *i.e.*, two ways of rendering such contacts through vibrations:

- *Impact (Impa)*: a 200 ms–long vibration burst is applied when the fingertip makes contact with the object; the amplitude of the vibration is proportional to the speed of the fingertip at the moment of the contact.
- *Distance (Dist)*: a continuous vibration is applied whenever the fingertip is in contact with the object; the amplitude of the vibration is proportional to the interpenetration between the fingertip and the virtual cube surface.

The implementation of these two techniques have been tuned according to the results of a preliminary experiment. Three participants were asked to carry out a series of push and grasp tasks similar to those used in the actual experiment. Results showed that 95 % of the contacts between the fingertip and the virtual cube happened at speeds below  $1.5 \text{ m s}^{-1}$ . We also measured the perceived minimum amplitude to be 15 % (0.6 g) of the maximum amplitude of the motors we used. For this reason, we designed the Impact vibration technique (Impa) so that contact speeds from  $0 \text{ m s}^{-1}$  to  $1.5 \text{ m s}^{-1}$  are linearly mapped into 15 % to 100 % amplitude commands for the motors. Similarly, we designed the distance vibration technique (Dist) so that interpenetrations from 0 cm to 2.5 cm are linearly mapped into 15 % to 100 % amplitude commands for the motors, recalling that the virtual cube has an edge of 5 cm.

### B. Experimental Design

We considered the same two tasks as in Experiment #1, described in Sec. III-B, that we analyzed separately, considering four independent, within-subject variables:

- *Vibrotactile Positioning*: the five positionings for providing vibrotactile hand rendering of the virtual contacts, as described in Sec. IV-A1.
- *Contact Vibration Technique*: the two contact vibration techniques, as described in Sec. IV-A2.
- *Visual Hand Rendering*: two visual hand renderings from the first experiment, Skeleton (Skel) and None, as

described in Sec. III-A; we considered Skeleton as it performed the best in terms of performance and perceived effectiveness and None as reference.

- *Target*: we considered target volumes located at NW and SW during the Push task, and at NE, NW, SW, and SE during the Grasp task (see Fig. 8); we considered these targets because they presented different difficulties.

To account for learning and fatigue effects, the positioning of the vibrotactile hand rendering (Positioning) was counter-balanced using a balanced  $10 \times 10$  Latin square. In these ten blocks, all possible Technique  $\times$  Hand  $\times$  Target combination conditions were repeated three times in a random order. As we did not find any relevant effect of the order in which the tasks were performed in the first experiment, we fixed the order of the tasks: first, the Push task and then the Grasp task.

This design led to a total of 5 vibrotactile positionings  $\times$  2 vibration contact techniques  $\times$  2 visual hand rendering  $\times$  (2 targets on the Push task + 4 targets on the Grasp task)  $\times$  3 repetitions = 420 trials per participant.

### C. Apparatus and Protocol

Apparatus and protocol were very similar to the first experiment, as described in Sec. III-D and Sec. III-E, respectively. We report here only the differences.

We employed the same vibrotactile device used by [50]. It is composed of two encapsulated Eccentric Rotating Mass (ERM) vibration motors (Pico-Vibe 304-116, Precision Microdrive, UK). They are small and very light (5 mm  $\times$  20 mm, 1.2 g) actuators capable of vibration frequencies from 120 Hz to 285 Hz and amplitudes from 0.2 g to 1.15 g. They have a latency of 20 ms that we partially compensated for at the software level with slightly larger colliders to trigger the vibrations very close the moment the finger touched the cube. These two outputs vary linearly together, based on the tension applied. They were controlled by an Arduino Pro Mini (3.3 V) and a custom board that delivered the tension independently to each motor. A small 400 mAh Li-ion battery allowed for 4 hours of constant vibration at maximum intensity. A Bluetooth module (RN42XV module, Microchip Technology Inc., USA) mounted on the Arduino ensured wireless communication with the HoloLens 2.

To ensure minimal encumbrance, we used the same two motors throughout the experiment, moving them to the considered positioning before each new block. Thin self-gripping straps were placed on the five positionings, with an elastic strap stitched on top to place the motor, as shown in Fig. 7. The straps were fixed during the entirety of the experiment to ensure similar hand tracking conditions. We confirmed that this setup ensured a good transmission of the rendering and guaranteed a good hand tracking performance, that was measured to be constant (15 ms) with and without motors, regardless their positioning. The control board was fastened to the arm with an elastic strap. Finally, participants wore headphones diffusing brown noise to mask the sound of the vibrotactile motors.

We improved the hand tracking performance of the system by placing on the table a black sheet that absorbs the infrared

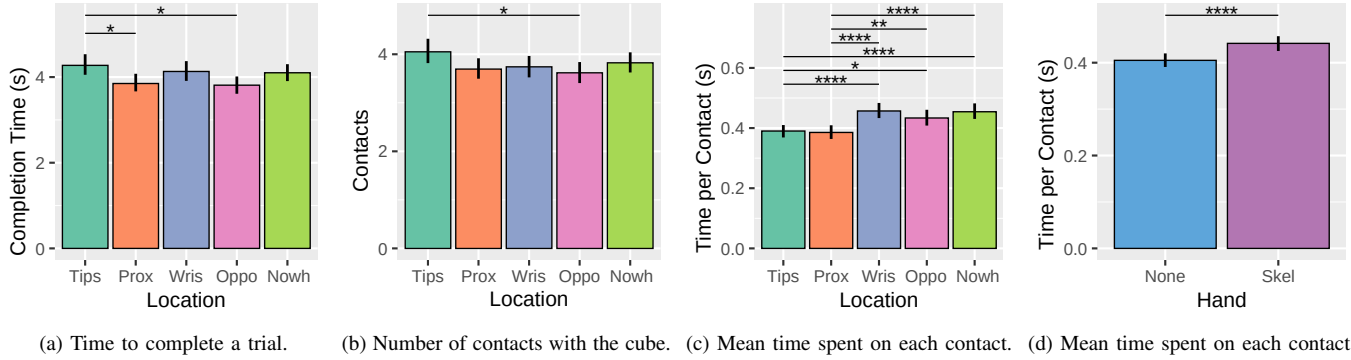


Figure 9. Experiment #2: Push task. Geometric means with bootstrap 95 % confidence interval for each vibrotactile positioning (a, b, and c) or visual hand rendering (d) and Tukey's HSD pairwise comparisons: \*\*\* is  $p < 0.001$ , \*\* is  $p < 0.01$ , and \* is  $p < 0.05$ .

light as well as placing the participants in front of a wall to ensure a more constant exposure to the light. We also made grasping easier by adding a grasping helper, similar to UltraLeap's Physics Hands.<sup>2</sup> When a phalanx collider of the tracked hand contacts the virtual cube, a spring with a low stiffness is created and attached between the cube and the collider. The spring pulls gently the cube toward the phalanxes in contact with the object so as to help maintain a natural and stable grasp. When the contact is lost, the spring is destroyed. Preliminary tests confirmed this approach.

#### D. Collected Data

During the experiment, we collected the same data as in the first experiment, see Sec. III-G.

At the end of the experiment, participants were asked if they recognized the different contact vibration techniques. They then rated the ten combinations of Positioning  $\times$  Technique using a 7-item Likert scale (1=Not at all, 7=Extremely): (*Vibration Rating*) How much do you like each vibrotactile rendering? (*Workload*) How demanding or frustrating was each vibrotactile rendering? (*Usefulness*) How useful was each vibrotactile rendering? (*Realism*) How realistic was each vibrotactile rendering? Finally, they rated the ten combinations of Positioning  $\times$  Hand on a 7-item Likert scale (1=Not at all, 7=Extremely): (*Positioning  $\times$  Hand Rating*) How much do you like each combination of vibrotactile location for each visual hand rendering?

#### E. Participants

Twenty subjects participated in the study (mean age = 26.8, SD = 4.1; 19 males, 1 female). One was left-handed, while the other nineteen were right-handed. They all used their dominant hand during the trials. They all had a normal or corrected-to-normal vision. Thirteen subjects participated also in the previous experiment.

Participants rated their expertise ("I use it more than once a year") with VR, AR, and haptics in a pre-experiment questionnaire. There were twelve experienced with VR, eight experienced with AR, and ten experienced with haptics. VR

and haptics expertise were highly correlated ( $r = 0.9$ ), as well as AR and haptics expertise ( $r = 0.6$ ). Other expertise correlations were low ( $r < 0.35$ ).

#### F. Results

Results were analyzed similarly as for the first experiment (see Sec. III-H). The LMM were fitted with the order of the five vibrotactile positionings (Order), the vibrotactile positionings (Positioning), the visual hand rendering (Hand), the contact vibration techniques (Technique), and the target volume position (Target), and their interactions as fixed effects and Participant as random intercept.

##### 1) Push Task:

**Completion Time:** On the time to complete a trial, there were two statistically significant effects: Positioning ( $F_{4,1990} = 3.8$ ,  $p = 0.004$ , see Fig. 9a) and Target ( $F_{1,1990} = 3.9$ ,  $p = 0.05$ ). Fingertips was slower than Proximal (+11%,  $p = 0.01$ ) or Opposite (+12%,  $p = 0.03$ ). There was no evidence of an advantage of Proximal or Opposite on No Vibrations, nor a disadvantage of Fingertips on No Vibrations. Yet, there was a tendency of faster trials with Proximal and Opposite. The NW target volume was also faster than the SW ( $p = 0.05$ ).

**Contacts:** On the number of contacts, there was one statistically significant effect of Positioning ( $F_{4,1990} = 2.4$ ,  $p = 0.05$ , see Fig. 9b). More contacts were made with Fingertips than with Opposite (+12%,  $p = 0.03$ ). This could indicate more difficulties to adjust the virtual cube inside the target volume.

**Time per Contact:** On the mean time spent on each contact, there were two statistically significant effects of Positioning ( $F_{4,1990} = 11.5$ ,  $p < 0.001$ , see Fig. 9c) and of Hand ( $F_{1,1990} = 16.1$ ,  $p < 0.001$ , see Fig. 9d) but not of the Positioning  $\times$  Hand interaction. It was shorter with Fingertips than with Wrist (-15%,  $p < 0.001$ ), Opposite (-11%,  $p = 0.01$ ), or NoVi (-15%,  $p < 0.001$ ); and shorter with Proximal than with Wrist (-16%,  $p < 0.001$ ), Opposite (-12%,  $p = 0.005$ ), or No Vibrations (-16%,  $p < 0.001$ ). This showed different strategies to adjust the cube inside the target volume, with faster repeated pushes with the Fingertips and Proximal positionings. It was also shorter with None

<sup>2</sup> <https://docs.ultraLeap.com/unity-api/Preview/physics-hands.html>

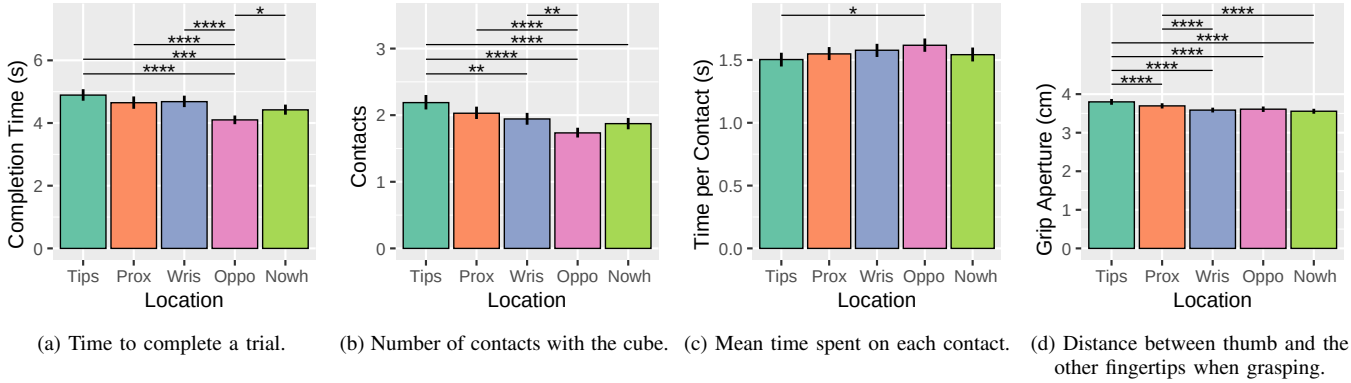


Figure 10. Experiment #2: Grasp task. Geometric means with bootstrap 95 % confidence interval for each vibrotactile positioning and Tukey’s HSD pairwise comparisons: \*\*\* is  $p < 0.001$ , \*\* is  $p < 0.01$ , and \* is  $p < 0.05$ .

than with Skeleton ( $-9\%$ ,  $p < 0.001$ ). This indicates, as for the first experiment, more confidence with a visual hand rendering.

## 2) Grasp Task:

**Completion Time:** On the time to complete a trial, there were two statistically significant effects: Positioning ( $F_{4,3990} = 13.6$ ,  $p < 0.001$ , see Fig. 10a) and Target ( $F_{3,3990} = 18.8$ ,  $p < 0.001$ ). Opposite was faster than Fingertips ( $+19\%$ ,  $p < 0.001$ ), Proximal ( $+13\%$ ,  $p < 0.001$ ), Wrist ( $+14\%$ ,  $p < 0.001$ ), and No Vibrations ( $+8\%$ ,  $p = 0.03$ ). No Vibrations was faster than Fingertips ( $+11\%$ ,  $p < 0.001$ ).

SE was faster than NE ( $p < 0.001$ ), NW ( $p < 0.001$ ), and SW ( $p < 0.001$ ); and SW was faster than NE ( $p = 0.03$ ).

**Contacts:** On the number of contacts, there were two statistically significant effects: Positioning ( $F_{4,3990} = 15.1$ ,  $p < 0.001$ , see Fig. 10b) and Target ( $F_{3,3990} = 7.6$ ,  $p < 0.001$ ). Fewer contacts were made with Opposite than with Fingertips ( $-26\%$ ,  $p < 0.001$ ), Proximal ( $-17\%$ ,  $p < 0.001$ ), or Wrist ( $-12\%$ ,  $p = 0.002$ ); but more with Fingertips than with Wrist ( $+13\%$ ,  $p = 0.002$ ) or No Vibrations ( $+17\%$ ,  $p < 0.001$ ). It was also easier on SW than on NE ( $p < 0.001$ ), NW ( $p = 0.006$ ), or SE ( $p = 0.03$ ).

**Time per Contact:** On the mean time spent on each contact, there were two statistically significant effects: Positioning ( $F_{4,3990} = 2.9$ ,  $p = 0.02$ , see Fig. 10c) and Target ( $F_{3,3990} = 62.6$ ,  $p < 0.001$ ). It was shorter with Fingertips than with Opposite ( $+7\%$ ,  $p = 0.01$ ). It was also shorter on SE than on NE, NW or SW ( $p < 0.001$ ); but longer on SW than on NE or NW ( $p < 0.001$ ).

**Grip Aperture:** On the average distance between the thumb’s fingertip and the other fingertips during grasping, there were two statistically significant effects: Positioning ( $F_{4,3990} = 30.1$ ,  $p < 0.001$ , see Fig. 10d) and Target ( $F_{3,3990} = 19.9$ ,  $p < 0.001$ ). It was longer with Fingertips than with Proximal ( $p < 0.001$ ), Wrist ( $p < 0.001$ ), Opposite ( $p < 0.001$ ), or No Vibrations ( $p < 0.001$ ); and longer with Proximal than with Wrist ( $p < 0.001$ ) or No Vibrations ( $p < 0.001$ ). But, it was shorter with NE than with NW or SW ( $p < 0.001$ ); and shorter with SE than with NW or SW ( $p < 0.001$ ).

3) **Discrimination of Vibration Techniques:** Seven participants were able to correctly discriminate between the two vibration techniques, which they described as the contact vibration (being the Impact technique) and the continuous vibration (being the Distance technique) respectively. Seven participants said they only felt differences of intensity with a weak one (being the Impact technique) and a strong one (being the Distance technique). Six participants did not notice the difference between the two vibration techniques. There was no evidence that the ability to discriminate the vibration techniques was correlated with the participants’ haptic or AR/VR expertise ( $r = 0.4$ ), nor that it had a statistically significant effect on the performance in the tasks. As the tasks had to be completed as quickly as possible, we hypothesize that little attention was devoted to the different vibration techniques. Indeed, some participants explained that the contact cues were sufficient to indicate whether the cube was being properly pushed or grasped. Although the Distance technique provided additional feedback on the interpenetration of the finger with the cube, it was not strictly necessary to manipulate the cube quickly.

4) **Questionnaire:** Fig. 11 shows the questionnaire results for each vibrotactile positioning. Questionnaire results were analyzed using Aligned Rank Transform (ART) non-parametric analysis of variance (see Sec. IV-D). Statistically significant effects were further analyzed with post-hoc pairwise comparisons with Holm-Bonferroni adjustment. Wilcoxon signed-rank tests were used for main effects and ART contrasts procedure for interaction effects. Only significant results are reported.

**Vibrotactile Rendering Rating:** There was a main effect of Positioning ( $F_{4,171} = 27.0$ ,  $p < 0.001$ ). Participants preferred Fingertips more than Wrist ( $p = 0.01$ ), Opposite ( $p < 0.001$ ), and No Vibration ( $p < 0.001$ ); Proximal more than Wrist ( $p = 0.007$ ), Opposite ( $p < 0.001$ ), and No Vibration ( $p < 0.001$ ); And Wrist more than Opposite ( $p = 0.01$ ) and No Vibration ( $p < 0.001$ ).

**Positioning  $\times$  Hand Rating:** There were two main effects of Positioning ( $F_{4,171} = 20.6$ ,  $p < 0.001$ ) and of Hand ( $F_{1,171} = 12.2$ ,  $p < 0.001$ ). Participants preferred Fingertips more than Wrist ( $p = 0.03$ ), Opposite ( $p < 0.001$ ), and No

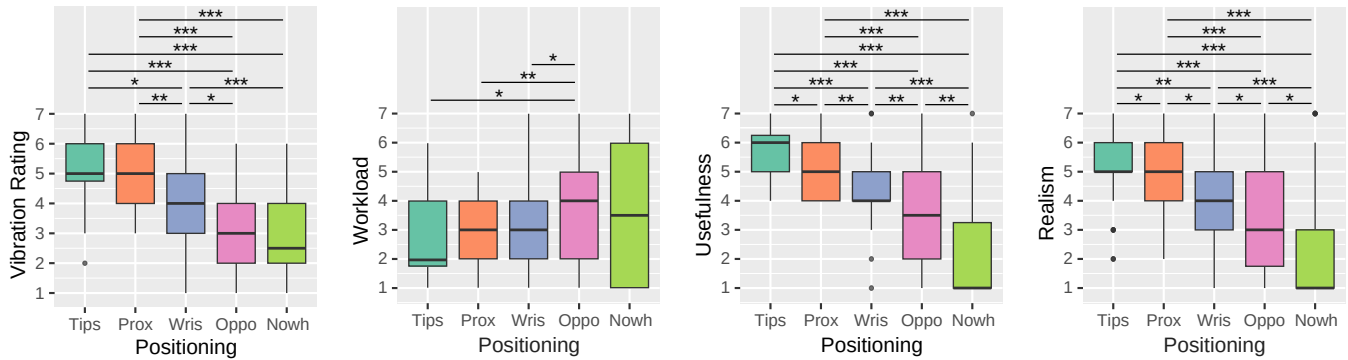


Figure 11. Experiment #2. Boxplots of the questionnaire results of each vibrotactile positioning and pairwise Wilcoxon signed-rank tests with Holm-Bonferroni adjustment: \*\*\* is  $p < 0.001$ , \*\* is  $p < 0.01$ , and \* is  $p < 0.05$ .

Vibration ( $p < 0.001$ ); Proximal more than Wrist ( $p = 0.003$ ), Opposite ( $p < 0.001$ ), and No Vibration ( $p < 0.001$ ); Wrist more than Opposite ( $p = 0.03$ ) and No Vibration ( $p < 0.001$ ); And Skeleton more than No Hand ( $p < 0.001$ ).

**Workload:** There was a main effect of Positioning ( $F_{4,171} = 3.9$ ,  $p = 0.004$ ). Participants found Opposite more fatiguing than Fingertips ( $p = 0.01$ ), Proximal ( $p = 0.003$ ), and Wrist ( $p = 0.02$ ).

**Usefulness:** There was a main effect of Positioning ( $F_{4,171} = 38.0$ ,  $p = 0.041$ ). Participants found Fingertips the most useful, more than Proximal ( $p = 0.02$ ), Wrist ( $p < 0.001$ ), Opposite ( $p < 0.001$ ), and No Vibrations ( $p < 0.001$ ); Proximal more than Wrist ( $p = 0.008$ ), Opposite ( $p < 0.001$ ), and No Vibrations ( $p < 0.001$ ); Wrist more than Opposite ( $p = 0.008$ ) and No Vibrations ( $p < 0.001$ ); And Opposite more than No Vibrations ( $p = 0.004$ ).

**Realism:** There was a main effect of Positioning ( $F_{4,171} = 28.8$ ,  $p < 0.001$ ). Participants found Fingertips the most realistic, more than Proximal ( $p = 0.05$ ), Wrist ( $p = 0.004$ ), Opposite ( $p < 0.001$ ), and No Vibrations ( $p < 0.001$ ); Proximal more than Wrist ( $p = 0.03$ ), Opposite ( $p < 0.001$ ), and No Vibrations ( $p < 0.001$ ); Wrist more than Opposite ( $p = 0.03$ ) and No Vibrations ( $p < 0.001$ ); And Opposite more than No Vibrations ( $p = 0.03$ ).

## G. Discussion

We evaluated sixteen visuo-haptic renderings of the hand, in the same two virtual object manipulation tasks in AR as in the first experiment, as the combination of two vibrotactile contact techniques provided at four delocalized positions on the hand with the two most representative visual hand renderings established in the first experiment.

In the Push task, vibrotactile haptic hand rendering has been proven beneficial with the Proximal positioning, which registered a low completion time, but detrimental with the Fingertips positioning, which performed worse (see Fig. 9a) than the Proximal and Opposite (on the contralateral hand) positionings. The cause might be the intensity of vibrations, which many participants found rather strong and possibly distracting when provided at the fingertips. This result was also

observed by Bermejo *et al.* [51], who provided vibrotactile cues when pressing a virtual keypad. Another reason could be the visual impairment caused by the vibrotactile motors when worn on the fingertips, which could have disturbed the visualization of the virtual cube.

We observed different strategies than in the first experiment for the two tasks. During the Push task, participants made more and shorter contacts to adjust the cube inside the target volume (see Fig. 9b and Fig. 9c). During the Grasp task, participants pressed the cube 25 % harder on average (see Fig. 10d). The Fingertips and Proximal positionings led to a slightly larger grip aperture than the others. We think that the proximity of the vibrotactile rendering to the point of contact made users to take more time to adjust their grip in a more realistic manner, *i.e.*, closer to the surface of the cube. This could also be the cause of the higher number of failed grasps or cube drops: indeed, we observed that the larger the grip aperture, the higher the number of contacts. Consequently, the Fingertips positioning was slower (see Fig. 10a) and more prone to error (see Fig. 10b) than the Opposite and Nowhere positionings.

In both tasks, the Opposite positioning also seemed to be faster (see Fig. 9a) than having no vibrotactile hand rendering (Nowhere positioning). However, participants also felt more workload (see Fig. 11) with this positioning opposite to the site of the interaction. This result might mean that participants focused more on learning to interpret these sensations, which led to better performance in the long run.

Overall, many participants appreciated the vibrotactile hand renderings, commenting that they made the tasks more realistic and easier. However, the closer to the contact point, the better the vibrotactile rendering was perceived (see Fig. 11). This seemed inversely correlated with the performance, except for the Nowhere positioning, *e.g.*, both the Fingertips and Proximal positionings were perceived as more effective, useful, and realistic than the other positionings despite lower performance.

Considering the two tasks, no clear difference in performance or appreciation was found between the two contact vibration techniques. While the majority of participants discriminated the two different techniques, only a minority identified them correctly (see Sec. IV-F3). It seemed that the

Impact technique was sufficient to provide contact information compared to the Distance technique, which provided additional feedback on interpenetration, as reported by participants.

No difference in performance was found between the two visual hand renderings, except for the Push task, where the Skeleton hand rendering resulted again in longer contacts. Additionally, the Skeleton rendering was appreciated and perceived as more effective than having no visual hand rendering, confirming the results of our first experiment. Participants reported that this visual hand rendering provided good feedback on the status of the hand tracking while being constrained to the cube, and helped with rotation adjustment in both tasks. However, many also felt that it was a bit redundant with the vibrotactile hand rendering. Indeed, receiving a vibrotactile hand rendering was found by participants as a more accurate and reliable information regarding the contact with the cube than simply seeing the cube and the visual hand reacting to the manipulation. This result suggests that providing a visual hand rendering may not be useful during the grasping phase, but may be beneficial prior to contact with the virtual object and during position and rotation adjustment, providing valuable information about the hand pose. It is also worth noting that the improved hand tracking and grasp helper improved the manipulation of the cube with respect to the first experiment, as shown by the shorter completion time during the Grasp task. This improvement could also be the reason for the smaller differences between the Skeleton and the None visual hand renderings in this second experiment.

In summary, the positioning of the vibrotactile haptic rendering of the hand affected on the performance and experience of users manipulating virtual objects with their bare hands in AR. The closer the vibrotactile hand rendering was to the point of contact, the better it was perceived in terms of effectiveness, usefulness, and realism. These subjective appreciations of wearable haptic hand rendering for manipulating virtual objects in AR were also observed by Maisto *et al.* [13] and Meli *et al.* [14]. However, the best performance was obtained with the farthest positioning on the contralateral hand, which is somewhat surprising. This apparent paradox could be explained in two ways. On the one hand, participants behave differently when the haptic rendering was given on the fingers, close to the contact point, with shorter pushes and larger grip apertures. This behavior has likely given them a better experience of the tasks and more confidence in their actions, as well as leading to a lower interpenetration/force applied to the cube [52]. On the other hand, the unfamiliarity of the contralateral hand positioning caused participants to spend more time understanding the haptic stimuli, which might have made them more focused on performing the task. In terms of the contact vibration technique, the continuous vibration technique on the finger interpenetration (Distance technique) did not make a difference to performance, although it provided more information. Participants felt that vibration bursts were sufficient (Impact technique) to confirm contact with the virtual object. Finally, it was interesting to note that the visual hand renderings was appreciated but felt less necessary when provided together with vibrotactile hand rendering, as the latter was deemed sufficient for acknowledging the contact.

As we already said in Sec. III-I, these results have some limitations as they address limited types of visuo-haptic renderings and manipulations were restricted to the thumb and index fingertips. While the simpler vibration technique (Impact technique) was sufficient to confirm contacts with the cube, richer vibrotactile renderings may be required for more complex interactions, such as collision or friction renderings between objects [52], [53] or texture rendering [54], [55]. More generally, a broader range of haptic sensations should be considered, such as pressure or stretching of the skin [13], [16]. However, moving the point of application of the sensation away may be challenging for some types of haptic rendering. Similarly, as the interactions were limited to the thumb and index fingertips, positioning a delocalized haptic rendering over a larger area of the hand could be challenging and remains to be explored. Also, given that some users found the vibration rendering too strong, adapting/personalizing the haptic feedback to one's preference (and body positioning) might also be a promising approach. Indeed, personalized haptics is recently gaining interest in the community [56], [57].

## V. CONCLUSION

This paper presented two human subject studies aimed at better understanding the role of visuo-haptic rendering of the hand during virtual object manipulation in OST-AR. The first experiment compared six visual hand renderings in two representative manipulation tasks in AR, *i.e.*, push-and-slide and grasp-and-place of a virtual object. Results show that a visual hand rendering improved the performance, perceived effectiveness, and user confidence. A skeleton rendering, providing a detailed view of the tracked joints and phalanges while not hiding the real hand, was the most performant and effective. The second experiment compared, in the same two manipulation tasks as before, sixteen visuo-haptic renderings of the hand as the combination of two vibrotactile contact techniques, provided at four different delocalized positions on the hand, and with the two most representative visual hand renderings established in the first experiment, *i.e.*, the skeleton hand rendering and no hand rendering. Results show that delocalized vibrotactile haptic hand rendering improved the perceived effectiveness, realism, and usefulness when it is provided close to the contact point. However, the farthest positioning on the contralateral hand gave the best performance even though it was disliked: the unfamiliarity of the positioning probably caused the participants to take more effort to consider the haptic stimuli and to focus more on the task. The visual hand rendering was perceived less necessary than the vibrotactile haptic hand rendering, but still provided a useful feedback on the hand tracking.

Future work will focus on including richer types of haptic feedback, such as pressure and skin stretch, analyzing the best compromise between well-round haptic feedback and wearability of the system with respect to AR constraints. As delocalizing haptic feedback seems to be a simple but very promising approach for haptic-enabled AR, we will keep including this dimension in our future study, even when considering other types of haptic sensations.

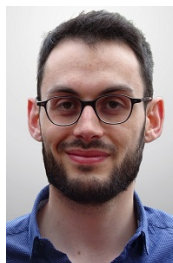
## ACKNOWLEDGMENT

This project has been funded by the European Union (ERC, ADVHANDTURE, 101088708).

## REFERENCES

- [1] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev, *3D User Interfaces: Theory and Practice*, 2017.
- [2] K. Kim, M. Billinghurst, G. Bruder, H. B.-L. Duh, and G. F. Welch, “Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008-2017)”, *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 11, pp. 2947–2962, 2018.
- [3] R. Xiao, J. Schwarz, N. Throm, A. D. Wilson, and H. Benko, “MRTouch: Adding Touch Input to Head-Mounted Mixed Reality”, *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 4, pp. 1653–1660, 2018.
- [4] M. Höll, M. Oberweger, C. Arth, and V. Lepetit, “Efficient Physics-Based Implementation for Realistic Hand-Object Interaction in Virtual Reality”, in *IEEE Virtual Real.*, 2018, pp. 175–182.
- [5] T. Piumsomboon, D. Altimira, H. Kim, A. Clark, G. Lee, and M. Billinghurst, “Grasp-Shell vs gesture-speech: A comparison of direct and indirect natural interaction techniques in augmented reality”, in *IEEE Int. Symp. Mix. Augment. Real.*, 2014, pp. 73–82.
- [6] M. Krichenbauer, G. Yamamoto, T. Taketom, C. Sandor, and H. Kato, “Augmented Reality versus Virtual Reality for 3D Object Manipulation”, *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 2, pp. 1038–1048, 2018.
- [7] M. Prachyabrued and C. W. Borst, “Visual feedback for virtual grasping”, in *IEEE Symp. 3D User Interfaces*, 2014, pp. 19–26.
- [8] T. Piumsomboon, A. Clark, M. Billinghurst, and A. Cockburn, “User-Defined Gestures for Augmented Reality”, in *Hum.-Comput. Interact.*, 2013, pp. 282–299.
- [9] T. Ha, S. Feiner, and W. Woo, “WeARHand: Head-worn, RGB-D camera-based, bare-hand user interface with visually enhanced depth perception”, in *IEEE Int. Symp. Mix. Augment. Real.*, 2014, pp. 219–228.
- [10] M. C. F. Macedo and A. L. Apolinário, “Occlusion Handling in Augmented Reality: Past, Present and Future”, *IEEE Trans. Vis. Comput. Graph.*, vol. 29, no. 2, pp. 1590–1609, 2023.
- [11] C. Diaz, M. Walker, D. A. Szafir, and D. Szafir, “Designing for Depth Perceptions in Augmented Reality”, in *IEEE Int. Symp. Mix. Augment. Real.*, 2017, pp. 111–122.
- [12] E. Peillard, F. Argelaguet, J.-M. Normand, A. Lécuyer, and G. Moreau, “Studying Exocentric Distance Perception in Optical See-Through Augmented Reality”, in *IEEE Int. Symp. Mix. Augment. Real.*, 2019, pp. 115–122.
- [13] M. Maisto, C. Pacchierotti, F. Chinello, G. Salvietti, A. De Luca, and D. Prattichizzo, “Evaluation of Wearable Haptic Systems for the Fingers in Augmented Reality Applications”, *IEEE Trans. Haptics*, vol. 10, no. 4, pp. 511–522, 2017.
- [14] L. Meli, C. Pacchierotti, G. Salvietti, *et al.*, “Combining Wearable Finger Haptics and Augmented Reality: User Evaluation Using an External Camera and the Microsoft HoloLens”, *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 4297–4304, 2018.
- [15] P. Lopes, S. You, A. Ion, and P. Baudisch, “Adding Force Feedback to Mixed Reality Experiences and Games using Electrical Muscle Stimulation”, in *CHI Conf. Hum. Factors Comput. Syst.*, 2018, pp. 1–13.
- [16] S.-Y. Teng, P. Li, R. Nith, J. Fonseca, and P. Lopes, “Touch&Fold: A Foldable Haptic Actuator for Rendering Touch in Mixed Reality”, in *CHI Conf. Hum. Factors Comput. Syst.*, 2021, pp. 1–14.
- [17] M. Al-Kalbani, I. Williams, and M. Frutos-Pascual, “Analysis of Medium Wrap Freehand Virtual Object Grasping in Exocentric Mixed Reality”, in *IEEE Int. Symp. Mix. Augment. Real.*, 2016, pp. 84–93.
- [18] A.-D. Blaga, M. Frutos-Pascual, M. Al-Kalbani, and I. Williams, “Usability Analysis of an Off-the-Shelf Hand Posture Estimation Sensor for Freehand Physical Interaction in Egocentric Mixed Reality”, in *IEEE Int. Symp. Mix. Augment. Real.*, 2017, pp. 31–34.
- [19] B. Yoon, H.-i. Kim, S. Y. Oh, and W. Woo, “Evaluating Remote Virtual Hands Models on Social Presence in Hand-based 3D Remote Collaboration”, in *IEEE Int. Symp. Mix. Augment. Real.*, 2020, pp. 520–532.
- [20] D. Saito, N. Wake, K. Sasabuchi, H. Koike, and K. Ikeuchi, “Contact Web Status Presentation for Freehand Grasping in MR-based Robot-teaching”, in *ACM/IEEE Int. Conf. Hum.-Robot Interact.*, 2021, pp. 167–171.
- [21] F. Argelaguet, L. Hoyet, M. Trico, and A. Lécuyer, “The role of interaction in virtual embodiment: Effects of the virtual hand representation”, in *IEEE Virtual Real.*, 2016, pp. 3–10.
- [22] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson, “Effects of Hand Representations for Typing in Virtual Reality”, in *IEEE Virtual Real.*, 2018, pp. 151–158.
- [23] V. Schwind, L. Lin, M. Di Luca, S. Jörg, and J. Hillis, “Touch with foreign hands: The effect of virtual hand appearance on visual-haptic integration”, in *Proc. ACM Symp. Appl. Percept.*, 2018, pp. 1–8.
- [24] M. Van Veldhuizen and X. Yang, “The Effect of Semi-Transparent and Interpenetrable Hands on Object Manipulation in Virtual Reality”, *IEEE Access*, vol. 9, pp. 17 572–17 583, 2021.
- [25] M. Sarac, T. M. Huh, H. Choi, M. R. Cutkosky, M. D. Luca, and A. M. Okamura, “Perceived Intensities of Normal and Shear Skin Stimuli Using a Wearable Haptic Bracelet”, *IEEE Robot. Autom. Lett.*, vol. 7, no. 3, pp. 6099–6106, 2022.
- [26] C. Pacchierotti, G. Salvietti, I. Hussain, L. Meli, and D. Prattichizzo, “The hRing: A wearable haptic device to avoid occlusions in hand tracking”, in *IEEE Haptics Symp.*, 2016, pp. 134–139.
- [27] J. E. Palmer, M. Sarac, A. A. Garza, and A. M. Okamura, “Haptic Feedback Relocation from the Fingertips to the Wrist for Two-Finger Manipulation in Virtual Reality”, in *IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2022, pp. 628–633.
- [28] V. Buchmann, T. Nilsen, and M. Billinghurst, “Interaction with Partially Transparent Hands and Objects”, 2005.
- [29] D. Kahl, M. Ruble, and A. Krüger, “Investigation of Size Variations in Optical See-through Tangible Augmented Reality”, in *IEEE Int. Symp. Mix. Augment. Real.*, 2021, pp. 147–155.
- [30] G. Bianchi, B. Knoerlein, G. Szekely, and M. Harders, “High Precision Augmented Reality Haptics”, in *Proc. EuroHaptics*, vol. 6, 2006, pp. 169–178.
- [31] S. Jeon and S. Choi, “Haptic Augmented Reality: Taxonomy and an Example of Stiffness Modulation”, *Presence*, vol. 18, no. 5, pp. 387–408, 2009.
- [32] B. Knoerlein, M. Di Luca, and M. Harders, “Influence of visual and haptic delays on stiffness perception in augmented reality”, in *IEEE Int. Symp. Mix. Augment. Real.*, 2009, pp. 49–52.
- [33] Y. Lee, S. Lee, and D. Lee, “Wearable Haptic Device for Stiffness Rendering of Virtual Objects in Augmented Reality”, *Appl. Sci.*, vol. 11, no. 15, p. 6932, 2021.
- [34] A. Hettiarachchi and D. Wigdor, “Annexing Reality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality”, in *CHI Conf. Hum. Factors Comput. Syst.*, 2016, pp. 1957–1967.
- [35] X. de Tinguy, C. Pacchierotti, M. Marchal, and A. Lécuyer, “Enhancing the Stiffness Perception of Tangible Objects in Mixed Reality Using Wearable Haptics”, in *IEEE Virtual Real.*, 2018, pp. 81–90.

- [36] S. V. Salazar, C. Pacchierotti, X. de Tinguy, A. Maciel, and M. Marchal, "Altering the Stiffness, Friction, and Shape Perception of Tangible Objects in Virtual Reality Using Wearable Haptics", *IEEE Trans. Haptics*, vol. 13, no. 1, pp. 167–174, 2020.
- [37] E. Normand and M. J. McGuffin, "Enlarging a Smartphone with AR to Create a Handheld VESAD (Virtually Extended Screen-Aligned Display)", in *IEEE Int. Symp. Mix. Augment. Real.*, 2018, pp. 123–133.
- [38] E. Pezent, A. Israr, M. Samad, *et al.*, "Tasbi: Multisensory Squeeze and Vibrotactile Wrist Haptics for Augmented and Virtual Reality", in *IEEE World Haptics Conf.*, 2019, pp. 1–6.
- [39] P. Kourtesis, F. Argelaguet, S. Vizcay, M. Marchal, and C. Pacchierotti, "Electrotactile feedback applications for hand and arm interactions", *IEEE Trans. Haptics*, vol. 15, no. 3, pp. 479–496, 2022.
- [40] M. Marchal and C. Pacchierotti, "Virtual Reality and Haptics", in *Robotics Goes MOOC*, B. Siciliano, Ed., 2022.
- [41] H. Ando, J. Watanabe, M. Inami, M. Sugimoto, and T. Maeda, "A fingernail-mounted tactile display for augmented reality systems", *Electron. Comm. Jpn. Pt. II*, vol. 90, no. 4, pp. 56–65, 2007.
- [42] F. Chinello, M. Malvezzi, D. Prattichizzo, and C. Pacchierotti, "A Modular Wearable Finger Interface for Cutaneous and Kinesthetic Interaction: Control and Evaluation", *IEEE Trans. Ind. Electron.*, vol. 67, no. 1, pp. 706–716, 2020.
- [43] S. Suzuki, H. Suzuki, and M. Sato, "Grasping a virtual object with a bare hand", in *ACM SIGGRAPH Posters*, 2014.
- [44] L. Meli, S. Scheggi, C. Pacchierotti, and D. Prattichizzo, "Wearable haptics and hand tracking via an RGB-D camera for immersive tactile experiences", in *ACM SIGGRAPH Posters*, 2014.
- [45] H. J. Kang, J.-h. Shin, and K. Ponto, "A Comparative Analysis of 3D User Interaction: How to Move Virtual Objects in Mixed Reality", in *IEEE Virtual Real.*, 2020, pp. 275–284.
- [46] M. Chessa, G. Maiello, L. K. Klein, V. C. Paulun, and F. Solari, "Grasping objects in immersive Virtual Reality", in *IEEE Virtual Real.*, 2019, pp. 1749–1754.
- [47] J. Bergström, T.-S. Dalsgaard, J. Alexander, and K. Hornbæk, "How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies", in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2021, pp. 1–20.
- [48] C. W. Borst and A. P. Indugula, "A Spring Model for Whole-Hand Virtual Grasping", *Presence*, vol. 15, no. 1, pp. 47–61, 2006.
- [49] D. Prattichizzo, C. Pacchierotti, and G. Rosati, "Cutaneous Force Feedback as a Sensory Subtraction Technique in Haptics", *IEEE Trans. Haptics*, vol. 5, no. 4, pp. 289–300, 2012.
- [50] L. Devigne, M. Aggravi, M. Bivaud, *et al.*, "Power Wheelchair Navigation Assistance Using Wearable Vibrotactile Haptics", *IEEE Trans. Haptics*, vol. 13, no. 1, pp. 52–58, 2020.
- [51] C. Bermejo, L. H. Lee, P. Chojecki, D. Przewozny, and P. Hui, "Exploring Button Designs for Mid-air Interaction in Virtual Reality: A Hexa-metric Evaluation of Key Representations and Multi-modal Cues", *Proc. ACM Hum.-Comput. Interact.*, vol. 5, pp. 1–26, 2021.
- [52] C. Pacchierotti, L. Meli, F. Chinello, M. Malvezzi, and D. Prattichizzo, "Cutaneous haptic feedback to ensure the stability of robotic teleoperation systems", *Int. J. Robot. Res.*, vol. 34, no. 14, pp. 1773–1787, 2015.
- [53] K. Kuchenbecker, J. Fiene, and G. Niemeyer, "Improving contact realism through event-based haptic feedback", *IEEE Trans. Vis. Comput. Graph.*, vol. 12, no. 2, pp. 219–230, 2006.
- [54] H. Culbertson, J. J. López Delgado, and K. J. Kuchenbecker, "One hundred data-driven haptic texture models and open-source methods for rendering on 3D objects", in *IEEE Haptics Symp.*, 2014, pp. 319–325.
- [55] S. Asano, S. Okamoto, and Y. Yamada, "Vibrotactile Stimulation to Increase and Decrease Texture Roughness", *IEEE Trans. Hum.-Mach. Syst.*, vol. 45, no. 3, pp. 393–398, 2015.
- [56] M. Malvezzi, F. Chinello, D. Prattichizzo, and C. Pacchierotti, "Design of Personalized Wearable Haptic Interfaces to Account for Fingertip Size and shape", *IEEE Trans. Haptics*, vol. 14, no. 2, pp. 266–272, 2021.
- [57] M. Umair, C. Sas, N. Chalabianloo, and C. Ersoy, "Exploring Personalized Vibrotactile and Thermal Patterns for Affect Regulation", in *Des. Interact. Syst. Conf.*, 2021, pp. 891–906.



**Erwan Normand** is a PhD student at the University of Rennes, France, since 2021. He received the BS degree in computer engineering from the Université de Technologie de Compiègne, France, in 2016 and the MS degree in software engineering from the École de Technologie Supérieure, Canada. He was a research engineer at Inria Bordeaux, France, from 2018 to 2021.



**Claudio Pacchierotti** (SM'20) is a tenured researcher at CNRS-IRISA in Rennes, France, since 2016. He was previously a postdoctoral researcher at the Italian Institute of Technology, Genova, Italy. Pacchierotti earned his PhD at the University of Siena in 2014. He was Visiting Researcher in the Penn Haptics Group at University of Pennsylvania in 2014, the Dept. of Innovation in Mechanics and Management at University of Padua in 2013, the Institute for Biomedical Technology and Technical Medicine (MIRA) at University of Twente in 2014, and the Dept. Computer, Control and Management Engineering of the Sapienza University of Rome in 2022. Pacchierotti received the 2014 EuroHaptics Best PhD Thesis Award and the 2022 CNRS Bronze Medal. He is Senior Chair of the IEEE Technical Committee on Haptics, Co-Chair of the IEEE Technical Committee on Telerobotics, and Secretary of the Eurohaptics Society.



**Eric Marchand** received the PhD degree and the "Habilitation diriger des recherches" in computer science from the Université de Rennes 1, Rennes, France, in 1996 and 2004, respectively. He is currently a professor of computer science with the same University and a member of the Inria/IRISA Rainbow team. His research interests include robotics, visual servoing, real-time object tracking, and augmented reality. He has been an associate editor of the IEEE Transactions on Robotics (2010-2014), IEEE Robotics and Automation Letters (2015-2018) and guest co-editor of The International Journal of Robotics Research (IJRR) special issue on "Robot Vision" (april 2015). He was a senior editor of the IEEE Robotics and Automation Letters (2018-2023).



**Maud Marchal** is a Full Professor in Computer Science at Univ. Rennes, INSA/IRISA, France. She received her Habilitation from the University of Rennes 1 in 2014 and her PhD from University Joseph Fourier in Grenoble in 2006. Her research interests include haptics, virtual reality and 3D interaction, physics-based simulation. She is Associate Editor of IEEE Transactions on Haptics since 2021.