Visual Feedback for Virtual Grasping

Mores Prachyabrued*
Mahidol University

Christoph W. Borst†
University of Louisiana at Lafayette

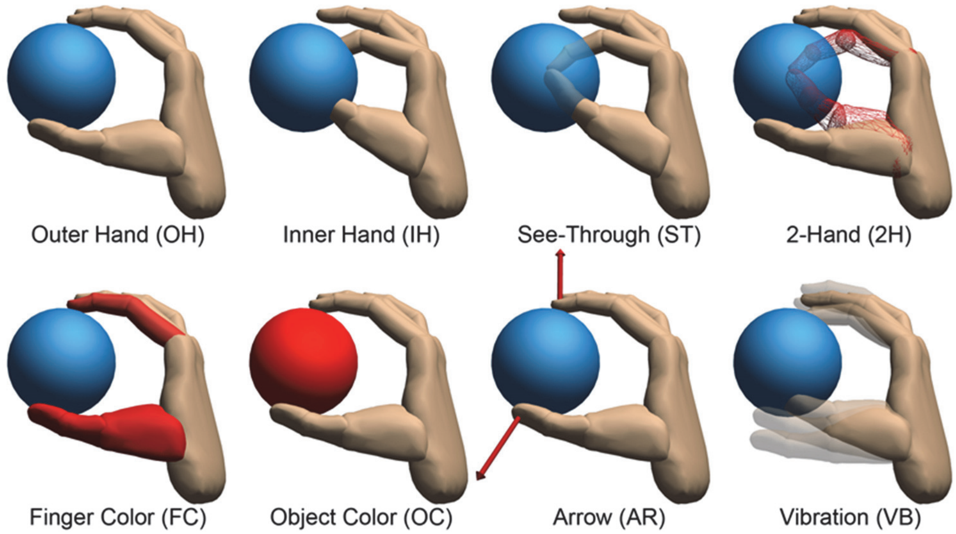

Figure 1: Visual feedback techniques investigated in our studies.

Abstract

We investigate visual feedback for virtual grasps, especially cues to improve behavior after real fingers enter a virtual object. To date, such visual cues have usually been developed in an ad-hoc manner, with minimal or no studies that can guide selection. Existing guidelines are based largely on other interaction types and provide inconsistent and potentially-misleading information when applied to grasping. We compare several different visual feedback types including those most commonly seen for virtual hand interaction and with some novel visual aspects. The visuals were tuned in a pilot study, and our main study evaluated results in terms of objective performance (finger penetration, release time, and precision) and subjective rankings. Performancewise, the most promising techniques all directly reveal penetrating hand configuration in some way. Subjectively, most techniques are better than simple interpenetrating visuals, with color changes being most promising. The results enable selection of the best cues based on the relevant tradeoffs. Results also provide a needed basis for more focused studies of specific visual cues and for better informing studies of multimodal feedback.

Keywords: Virtual grasping, visual feedback.

Index Terms: I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

*e-mail: mores_p@hotmail.com
†e-mail: cwborst@gmail.com

1 Introduction

Grasping quality is important in VR applications such as manual assembly training [1] or design review of vehicle control layouts [2]. There is also a growing interest in hand interaction for recent technologies: e.g., see [3, 4, 5, 6] for work using Kinect, Leap Motion, Digits, and interactive surfaces, respectively.

Hand-object interpenetration, where a hand sinks into virtual objects due to the lack of real physical constraints, is a fundamental problem for hand-based interaction. Past visual interpenetration studies, which usually did not consider grasping, are inconsistent and provide potentially-misleading guidelines if applied to grasping (Section 2). The relative effectiveness of different visual cues for grasping systems is also not known. We address this with a new study of visual feedback for grasping.

Interpenetration contributes to artifacts such as a "sticking object" when exaggerated finger motions are required for release, degrading performance and subjective experience [7]. Users may reduce such problems using "light touch" [8, 9]. Visual cues may help a user understand and control this light touch, as suggested by a prior study of two minimal (baseline) approaches [9]. Specifically, allowing visual interpenetration (Figure 1, inner hand, IH) produces lighter touch than a visually-constrained virtual hand (outer hand, OH), but users dislike IH. We seek to mitigate the tradeoff or provide better results than baselines.

Other approaches to deal with penetration include haptics to improve hand behavior [2, 10] and a special release mechanism [7] to reduce aftereffects of interpenetration. Even with such techniques, it is important to understand the impact of different visual approaches that may be used in combination with them, and to identify the best visual approaches for fair comparison to non-visual approaches. Visual rendering is almost always present with VR grasping, so we naturally should make good use of it.

We do not study haptic feedback here and we do not expect it to eliminate penetration in the near future. Whole-hand force devices are promising in some applications, but they have limits in degrees of freedom, stiffness, and earth-referenced forces without devices that are complex, costly, and restricting [11]. This may lead to more development of minimal or passive haptic approaches that can aid users but not constrain motion physically, e.g., [1, 2]. Visual feedback may be preferred to some approaches [1]. Additionally, recent hand sensing work increasingly points to optical hand tracking with minimal or no worn devices.

Regarding the approach of a heuristically-triggered release mechanism [7], visual cues to improve behavior remain important. Some grasp methods cannot readily incorporate the mechanism, e.g., [3], and most do not. Even when present, heuristics do not trigger consistently. Release motions influenced by mechanisms besides the real hand may have side effects proportionally to penetration depth [7, 12]. Light touch can also improve simulation stability and reduce visual-proprioceptive discrepancy for constrained visuals such as OH extensions.

We identify several possible visual techniques, considering those common in grasping systems, and also including novel aspects. Each technique was visually tuned by a pilot study and then a main study evaluated performance and subjective rankings. Two minimal techniques (IH and OH) provided baselines for comparisons. Section 3.2 gives the other technique motivations.

In summary, the top-level contributions of this paper are:

- Design and evaluation of visual cues to help users control grasping. Prior work on grasping cues is ad hoc, with little or no study guiding choices, and guidelines are inconsistent.
- Novel aspects in the visual cues: See-through, vibration, and a modified ghost hand (two-hand) techniques.
- Main result: The best performers reveal a penetrating hand directly. 2H gives good performance and reasonable user experience. Color techniques may be preferred subjectively with some performance compromise.

2 Related Work

Studies by Lindeman et al. [13] and Burns et al. [14] support the use of a constrained virtual hand rather than a penetrating hand, but the studies did not consider grasping. Such results support a standard 3D interaction guideline of avoiding penetrating visuals [15]. Several grasping systems included mechanisms to visually constrain a hand, e.g., [8, 16, 17].

In contrast, Prachyabrued and Borst [9] showed how a constrained hand reduces performance and causes users to misunderstand grasp. Durlach et al. [18] previously showed similar results for a pointing task. Our findings in the following sections show that, performancewise, the penetrating and constrained hand baselines rank among the best and worst of several techniques. Furthermore, we show how some techniques mitigate the tradeoffs between these baselines.

Lindeman et al. also considered changing finger color to communicate penetration depth to improve users' understanding [13]. Color changes have been included in various grasping systems, e.g., [1, 2, 16, 19]. Ullmann and Sauer [19] suggest that a ghost hand technique would be helpful in combination with discrete color effects. They changed phalanx color to indicate contact and changed whole hand color to indicate a valid grasp. The ghost technique represented real hand in wireframe while also showing a constrained solid hand. Gomes de Sá and Zachmann [1] used discrete object color changes and a ghost technique for grasped objects, and they also considered vibrotactile cues. Their results suggest the object color effects are preferred over vibrotactile collision cues. Moehring and Froehlich [2] showed that discrete finger segment color changes, indicating segments defining a grasp, can improve subjective ratings.

Rusak et al. [20] made an object transparent to show contact region, through an object, for a constrained hand. This improved finger positioning on a block. Unlike this work, we present a transparency cue (ST) revealing a penetrating hand, and the object remains opaque with respect to other objects besides the hand.

Fabiani et al. [10] studied visual and auditory cues as a substitute for, and in combination with, force feedback. The visual cues consisted of LED-type force level indicators. The grasped object was deformable, avoiding the interpenetration problem. The study asked subjects to compress the object by a certain amount while moving to a target. The main result was that all feedback types reduced the amount by which the object was squeezed, and force feedback provided a more even squeezing across fingers. In contrast, we consider the interpenetration problem, we study several more direct visual feedback types, we use a pick-and drop task without specific instructions about hand closure, and we include subjective results.

Hand visual fidelity has been studied, e.g., [18, 21, 22]. A moderately-realistic 3D hand model has been seen to provide better targeting than abstract models [21] or very crude models [18]. In our study, we focus on visual cues added to a similarly-realistic 3D hand model, and on the grasping problem.

Visual cues for inter-object contact have been studied outside of grasping. The most relevant work, by Sreng et al. [23], designed proximity and contact cues for assembly or maintenance simulation. They included a lighting effect illuminating contact areas and visual glyphs (arrow, disk, and sphere) conveying proximity, contact location, and contact force. Glyph properties, such as size, color, or deformation could be functions of proximity or force. Evaluation showed that subjects preferred color-coded glyphs for proximity, deformation for force, and lighting to mark contact area. In our work, we include per-finger arrow glyphs that scale with contact to show force or penetration. They resemble the force arrow glyphs for grasping shown by Borst and Indugula [8].

Several grasp techniques could be impacted by our findings because they share the finger penetration and release problem, e.g., [2, 3, 8, 24, 25]. A more detailed summary can be found in [7]. As already noted, some of these techniques include visual contact cues, but their relative effectiveness is not known.

3 Grasping Method and Visual Cue Design

3.1 Grasping Implementation

We implemented grasping using a virtual spring coupling between virtual and real (tracked) hands. The approach is known from previous work on physically-based grasping [8, 9]. The virtual hand model, also called the spring hand, is moved by a physics engine as a result of spring forces in the coupling. It is also affected by the physics engine's collision and response mechanisms, such that the spring hand remains outside object boundaries. The coupling contains one linear spring to pull the virtual palm towards the real palm, one torsional palm spring for palm rotation, and 20 torsional springs for finger joints. These joints follow a common model [26] with 3 joints per finger: a 2-dof metacarpophalangeal joint (MPJ) for first knuckle abduction and flexion, a 1-dof proximal interphalangeal joint (PIJ) for second knuckle flexion, and a distal interphalangeal joint (DIJ) for third knuckle flexion. The thumb has a 2-dof trapeziometacarpal joint (TMJ) in the palm for roll and abduction, a 1-dof MPJ for first knuckle flexion, and a 1-dof IJ for second knuckle flexion.

The visual hand geometry is derived from a Viewpoint Datalabs model with 16 rigid segments. The physics engine is the NVIDIA PhysX SDK. Each of the 20 finger springs is implemented using a PhysX revolute joint. The palm springs are implemented using equations from [8]. The thumb springs are 2.1 times stiffer than other digit springs, as suggested by prior work on optimizing relative spring stiffness [27]. PhysX is set to allow collision shapes to overlap by 0.6 cm for improved contact simulation. Correspondingly, hand collision shape is larger than visual shape.

3.2 Visual Feedback Techniques and Parameters

We describe the visual feedback techniques (Figure 1) and summarize results of a pilot study that tuned these visuals. Due to space constraints, we do not detail the pilot study beyond basic results. The purpose of the study was to identify reasonable parameters and gain preliminary insight into the techniques. The apparatus and virtual environment can be seen in an accompanying video or in Section 4. Five subjects (two VR experts and three novices) adjusted parameters for visual cues with instructions to "encourage light touch". They also specified the strength of their preferences for each parameter. Subjects handled a 5.25 cm-radius ball and were first introduced to visual penetration, the sticking problem, and light touch. Two methods, IH and OH, were not included because they are not adjustable.

3.2.1 Inner Hand (IH)

Inner Hand is a baseline technique that presents a 3D hand model following the user's real (tracked) hand. It produces unpleasant visual interpenetration, but it is known to produce lighter touch for better release performance when compared to Outer Hand [9].

3.2.2 Outer Hand (OH)

Outer Hand is a baseline technique presenting a 3D hand model constrained to avoid visual interpenetration. We implement it by showing the simulation-controlled spring hand configuration (Section 3.1). Users report OH as more natural than IH in [9].

3.2.3 See-Through (ST)

See-Through shows an inner hand model but reveals the penetrating portion, motivated by the potential to have the better performance of IH while making the visual interpenetration less disturbing. To our knowledge, this is a novel rendering style for grasping (Section 2). It somewhat resembles augmented reality (AR) systems where real hand parts are seen, e.g., [3], so results may give some insight into acceptability of this AR feature.

ST makes front-facing surfaces of grasped objects appear semitransparent at pixels that would otherwise occlude the inner hand. Our OpenGL implementation is summarized as:

1. Clear stencil buffer values.
2. Render object's back faces to depth buffer.
3. Render hand while setting stencil values of hand pixels (passing depth buffer test).
4. Render object's front faces semitransparently (alpha < 1), using stencil buffer to draw only on penetrating hand pixels.
5. Render object's front faces solidly, using stencil buffer to avoid drawing over penetrating hand pixels.

Based on our pilot study, the transparency level was tuned to alpha of 0.65 (mean best value, normalized alpha range).

3.2.4 2-Hand (2H)

2-Hand resembles ghost hand techniques and shows both inner and outer hand models. 2H attempts to combine the natural impression of OH with additional understanding about real hand state provided by IH, which may improve hand control. Both ST and 2H can be considered ways to directly reveal inner hand state.

We consider 12 presentation styles for 2H (Figure 2). Users can adjust three parameters called ghost position, rendering type, and inner hand visibility. Ghost position selects which of the two hands is drawn differently than normal, thereby specifying whether the inner or outer hand is considered the ghost (Figure 2a vs. 2d). Rendering types for the ghost hand are colored-semitransparent (Figure 2a), wireframe (Figure 2b), and skin-semitransparent (Figure 2c). The visibility options are see-through or occluded, referring to the appearance of the inner hand with respect to the grasped object (Figure 2b vs 2e). The see-through transparency is adjustable, but we re-use ST level for consistency.

Previous work mainly uses an occluded inner ghost with wireframe or colored-semitransparent rendering. We added the see-through option to combine ghost techniques with ST.

Based on the pilot study, we chose the best parameters of inner position, wireframe rendering, and see-through visibility.

3.2.5 Finger Color (FC)

Finger Color presents an outer hand model with fingers colored based on grasp. As finger closure increases, color changes continuously from normal to red. Unlike the above techniques, FC and the remaining techniques represent closure indirectly.

We use red because we expect it has strong visual impact, suggests "stop", and because it is common in prior work, e.g., [1, 8, 13, 23]. We consider different rates and mapping (interpolation) types for changing color from normal to red, and we also consider penetration vs. force representation. Mapping types are linear, superlinear, and sublinear. Nonlinear mapping may be useful to increase impact in a certain part of motion range or to counteract perceptual nonlinearities in color change. Penetration magnitude, p, is computed per digit as the distance between its tips in outer and inner hand models. Normalized penetration, p_n, is $\min(p/R, 1)$, where R = 5.25 cm, the radius of the ball in the pilot study. The mapping function converts p_n to normalized feedback intensity, i_n. Linear mapping is $i_n = p_n$, sublinear is $i_n = p_n^{1.5}$, and superlinear is $i_n = 1\text{-}(1\text{-}\ p_n)^{1.5}$. Rate is set by a scale factor, s_f. Per penetrating finger, we then compute an RGB vector for color, fc, as:

$$fc = SC + (RC - SC) \cdot s_f \cdot i_n\,, \qquad (1)$$

where SC is normal color and RC is red (RGB vectors). Color channels are clamped so change does not exceed the RC target.

We also consider two closure representations: penetration and force. Penetration representation is as described above. For force representation, a feature of virtual springs is incorporated so color better represents forces exerted on the object. Due to the thumb spring scaling (Section 3.1), the thumb penetrates less than a finger when exerting equal force (however, when multiple fingers are used, the thumb exerts more force than one finger). The force representation multiplies normalized feedback intensities for thumb and other digits by $m{\cdot}s$ and s, respectively, where m is thumb stiffness scale (here, 2.1) and s is $2/(m+1)$. This shifts some color to the thumb while maintaining an intensity sum, making color distribution match force distribution.

Based on the pilot study, linear mapping was chosen (preferred by 3 subjects), scale was set to average tuned value 1.0, and representation was set to penetration (preferred by all subjects).

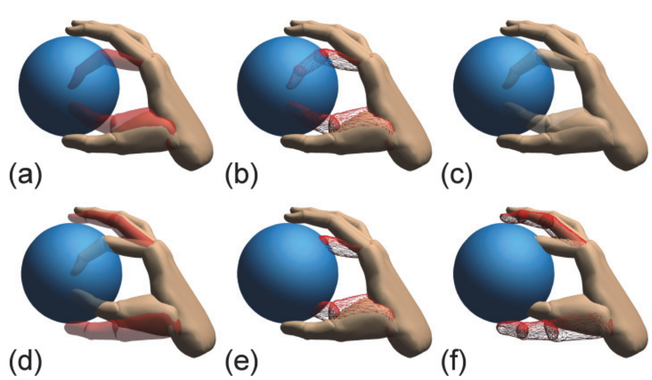

Figure 2: Some of the 2H variations: (a) colored-semitransparent inner see-through, (b) wireframe, (c) skin-semitransparent, (d) outer position, (e) occluded, (f) outer wireframe, occluded.

3.2.6 Object Color (OC)

Object Color presents an outer hand model and varies the grasped object's color. The color changes continuously from normal to red as closure increases. Since users are likely to focus more on grasped objects than on finger details [7], it is interesting to consider object color as an alternative to finger color.

As with finger color, we consider rate (scale), interpolation type, and closure representation. The calculations are analogous to those for finger color. However, only the digit giving maximum color intensity is used to determine object color, since there is only one object being colored.

Based on the pilot study, we chose linear mapping, scale of 1.3, and penetration representation.

3.2.7 Arrow (AR)

In the Arrow technique, arrow glyphs emerge from fingernails of an outer hand model, growing with increasing hand closure. We include AR to consider glyph approaches and because it was seen in the work on which our grasping implementation is based [8].

We again include adjustable rate (scale), interpolation type, and closure representation. Additionally, two arrow directions are considered: normal to fingernails or parallel to penetration vectors, where the penetration vectors point between outer hand digit tips and corresponding inner hand tips. The difference is shown in Figure 3. Arrow length is computed per digit as $s_a \cdot i_n \cdot R$, where s_a is scale and the other terms are as in Section 3.2.5.

From the pilot study, we chose tuned scale 1.06, penetration representation, and penetration-parallel arrows. Subjects indicated no strong mapping preference; we use linear to match FC and OC.

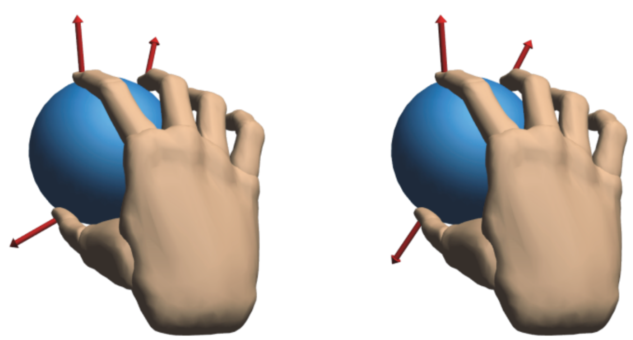

Figure 3: Arrow directions: normal (left) and penetration (right).

3.2.8 Vibration (VB)

Vibration is a novel technique that visually vibrates segments of an outer hand model. This affects visual hand appearance only and does not affect the grasped object or simulation. The vibration might be disturbing for some users, but, considering positive effects of visually-disturbing IH [9], this may discourage closure, inducing light touch for grasp training.

We consider variations of VB generated by different joint angles, amplitude- vs. frequency-based closure representation, and scale parameters (s_{va} for amplitude, s_{vf} for frequency).

Per contacting digit, we visually offset one joint angle by adding angle $r(t)$ for rendering:

$$\text{Amplitude-based:} \quad r(t) = i_n \cdot s_{va} \cdot \sin(s_{vf} \cdot 2\pi \cdot t) \,, \qquad (2a)$$

$$\text{Frequency-based:} \quad r(t) = s_{va} \cdot \sin(i_n \cdot s_{vf} \cdot 2\pi \cdot t) \,, \qquad (2b)$$

where t is time elapsed (sec) since finger contact began.

Based on the pilot study, we chose MPJ-flexion finger joint, TMJ-abduction thumb joint, and amplitude-based vibration with $s_{va} = 12.6$ and $s_{vf} = 9.04$.

4 Main Evaluation of Techniques

4.1 Design

We conducted within-subjects experiments to evaluate the 8 techniques. The first session was a targeted ball-drop experiment to measure finger penetration, release time, and translation error (release-imparted motion). The independent variables were:

1. *Visual technique* – OH, IH, ST, 2H, FC, OC, AR, and VB.
2. *Ball size* – small (4 cm radius) and large (6.5 cm).

Next, a technique ranking experiment asked subjects to rank the techniques based on visual appearance, behavior effect, and overall preference, while handling a medium ball (5.25 cm).

Subjects also reviewed and commented on techniques. We only mention subject comments briefly, where most relevant.

4.2 Subjects

30 subjects participated: 28 males and 2 females, aged 16 to 40 years (average 26), 28 right-handed and 2 left-handed. 25 subjects were students: 19 from computer-related fields. 5 non-students were also in computer-related fields.

4.3 Apparatus

Subjects reached into the mirror-based 3D display seen in Figure 4. An Acer GD235HZ LCD showed 1920 x 1080 images at 120 Hz, viewed with NVIDIA 3DVision glasses. A mirror with a polarizing layer reflected the LCD images. The head was not tracked – we noticed only minimal head motions when developing experiment tasks. An 18-sensor Cyberglove tracked finger joints, with missing distal joint angles computed as two thirds of middle knuckle angles. An Ascension miniBird 500 tracked the palm. Desktop speakers played audio feedback. Subjects ranked visual techniques using a Griffin PowerMate knob. The PC was a Dell Precision T5500 with a Xeon W3680 3.33GHz processor, 12 GB Ram, and an NVIDIA Quadro 5000 graphics card.

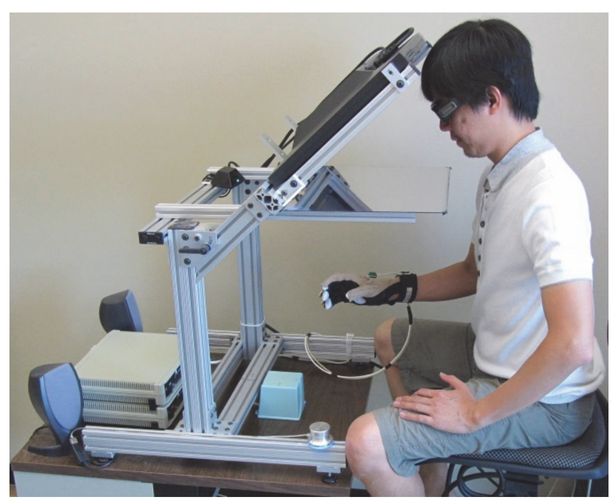

Figure 4: Equipment for our studies.

4.4 Procedure

We calibrated the CyberGlove per subject. Subjects then practiced OH-type grasping and releasing for 3 ball sizes in a simple ball-drop environment (similar to Figure 5, but without targeting). Subjects could use any digits or grasp shapes they found suitable.

4.4.1 Targeted Ball-Drop Procedure

This experiment had eight trial sets: one per visual technique, in random order. Per set, subjects first practiced grasping with the visual technique in the simple ball-drop environment. Subjects then practiced a targeted ball-drop task twice: once with on-screen instructions and once without. The instruction asked subjects to pick up the ball from the floor, move it into a wireframe cube above an X target, and release it on expiration of an audible countdown timer (Figure 5). Instructions also requested "NORMAL finger release motion (not too fast or too slow)". The cube switched from black to bright green as the ball was centered in it. The color switched at a threshold distance of 1.5 cm, which also triggered the 2-s countdown sound (tick-tick-beep). The ball center was required to stay within 1.75 cm of the cube center during the countdown or the trial restarted. Premature release (between pickup and beep sound) also restarted the trial.

Per technique set, after practice trials, subjects performed 8 real trials: 4 trials for each of the 2 ball sizes, in randomized order.

4.4.2 Technique Ranking Procedure

Subjects performed 3 ranking tasks, ordering icons representing techniques from "MOST" to "LEAST" for an on-screen question (Figure 6) while freely switching between techniques to test them. The first two tasks asked "how much you liked their visual appearances" and "how strongly they affected your behavior", in random order. The third task asked "how much you preferred them overall". Initial icon order was random per question.

Subjects rotated the knob to highlight any icon and test the highlighted visual technique. Subjects changed the position of an icon by pressing the knob while rotating. Subjects signaled the experimenter when done ranking.

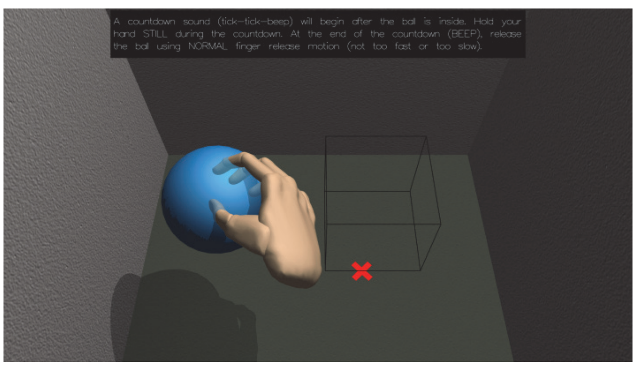

Figure 5: Targeted ball-drop task with wireframe cube and target.

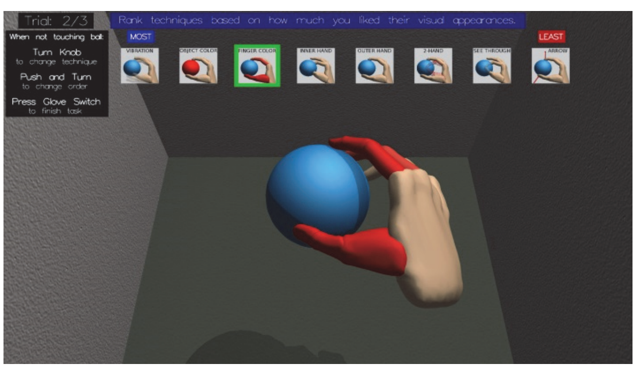

Figure 6: Technique ranking. Subjects ordered icons representing techniques while freely highlighting icons to test techniques.

5 Results and Discussion

We compare techniques to OH and IH. We want to identify which technique is most consistently found better than baselines, or at least provides the most promising tradeoff. Indirectly, this also gives insight into relative performance of techniques.

5.1 Results and Discussion for Targeted Ball-Drop

Performance, averaged over 4 trials per case, was computed as:

1. *Penetration depth*: maximum of finger penetrations at the end of the countdown, where per-finger penetration is the distance between the tracked finger tip and spring hand tip.
2. *Release time*: time between the countdown end (release start) and when no spring hand fingers touch the ball (release end).
3. *Translation error*: horizontal motion magnitude (translation parallel to floor) of ball from release start to floor contact [7].

We consider penetration depth to be the main indicator of behavioral impact. It was the most sensitive metric, with lowest within-technique dispersion. Penetration is a likely contributor to release effects [7, 9] and relates directly to the desired light touch.

Figures 7-9 and Table 1 summarize results. Tables reflect pooled ball sizes, as both sizes gave similar overall patterns of technique effects. Statistical tests are nonparametric because distributions tended to be positively skewed. We used Friedman tests for overall effects and a protected least-significant difference approach (PLSD) to follow-up testing with Wilcoxon signed-rank tests. PLSD is not conservative, but this can be mitigated by noting patterns of effects rather than isolated findings, particularly in borderline cases. We treat p-values below 0.05 as significant and refer to other values below 0.1 as showing "trends". Pragmatic readers should note that slightly higher p-values suggest likely, but unproven, effects, while conservative readers may prefer to discount borderline values. Final recommendations (Section 5.3.2) comment briefly on confidence.

Inspection of plots suggests that IH and OH provided best and worst overall performance, respectively. Considering statistical analyses, we can further categorize each technique into one of three groups based on differences from IH and OH:

1. Promising: Techniques showing the most potential: Better than OH, based on detected differences or trends, and no statistically-detected significance or trend compared to IH. Although differences from IH may be detected by more extensive experiments, they are not likely large or consistent.
2. Compromise: significance or trends of worse performance than IH, but better performance than OH.
3. Unpromising: no detected significance or trends compared to OH, but difference or trend of worse performance than IH.

Penetration depth: Visual technique impacted penetration depth, $\chi^2(7) = 40.0$, $p < .001$. Techniques grouped as follows:

- Promising: IH, ST, 2H.
- Compromise: OC, VB, AR, FC.
- Unpromising: OH.

Penetration was larger for the large ball (median 2.8 cm) than the small ball (1.8 cm), $\chi^2(1) = 30.0$, $p < .001$. We only give overall size effect, as per-visual testing does not add more insight.

Release time: Visual technique impacted the release time, $\chi^2(7) = 14.8$, $p = .039$. Techniques grouped as follows:

- Promising: IH, VB, 2H, AR.
- Compromise: OC and ST.
- Unpromising: FC and OH.

It took longer to release the large ball (median 0.56 s) than the small ball (0.4 s), $\chi^2(1) = 30.0$, $p < .001$.

Translation error: Visual technique impacted translation error, $\chi^2(7) = 19.2$, $p = .007$. Techniques grouped as follows:

- Promising: ST, IH, FC, 2H.
- Compromise: OC.
- Unpromising: VB, AR, OH.

Release was less accurate with the large ball (median error 0.98 cm) than with the small ball (0.7 cm), $\chi^2(1) = 13.3$, $p < .001$.

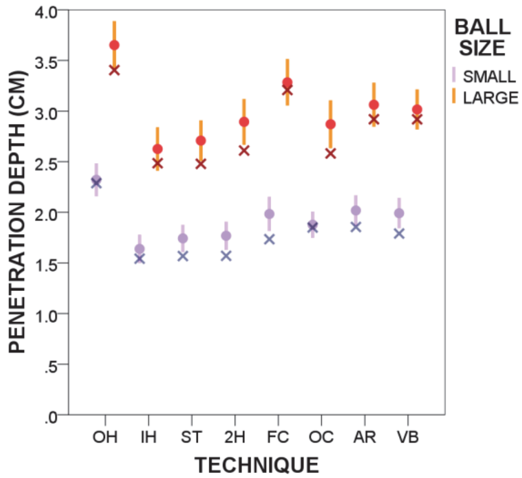

Figure 7: Penetration depth: medians (X marks), means (circles), and standard errors (bars centered on means).

Penetration was moderately correlated to release time ($r_s(1920) = .481$, $p < .001$) and translation error ($r_s(1920) = .443$, $p < .001$).

Discussion: Less required finger extension, due to less penetration, tends to provide faster release and less sticking, which also improves release accuracy. Where time and accuracy did not exactly follow penetration, this may reflect their reduced sensitivity: reaction time differences blur release time and palm motion at release blurs translation error. Some techniques could have additional impacts on release behavior that explains mixed results of VB, AR, or FC. Although additional experimentation can resolve this, its value is limited – these techniques did not have strong performance overall.

We give an overall summary by counting how often techniques appeared in categories above. The resulting list is also close to a penetration-only ranking. From most to least promising:

1. IH and 2H: Consistently appeared in "promising" category.
2. ST: In "promising" twice and in "compromise" once.
3. OC, VB, AR, FC: On average, in "compromise" category.
4. OH: Consistently in "unpromising" category.

The most promising techniques (IH, 2H, ST) directly revealed real hand configuration rather than using indirect representations of finger penetration. Results generalize the value of showing an inner hand over augmenting an outer hand, and they extend prior IH vs. OH findings [9] to show that IH and OH rank among the best and worst of several techniques. We found time differences not seen in [9], and showed performance of several techniques.

Ball size: larger finger penetration for the large ball may be due to the larger range of motion available or tighter grasps learned for larger objects that are expected to be heavier based on real-world experiences. Reduced performance with the large ball follows from increased penetration as discussed previously for visual technique. Results directly show the relationship between finger penetration and object size, speculated in [7].

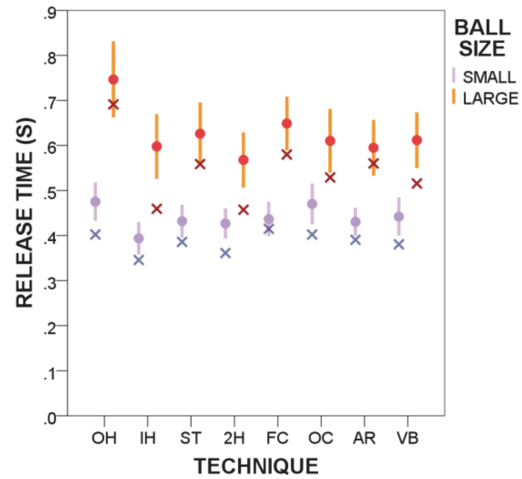

Figure 8: Release time: medians, means, and standard errors.

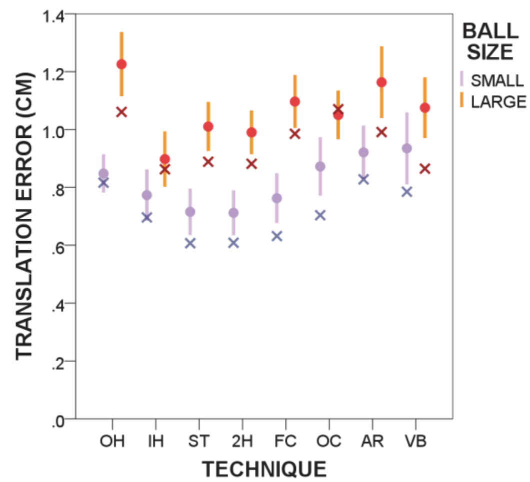

Figure 9: Translation error: medians, means, and standard errors.

5.2 Results and Discussion for Technique Ranking

Per subject and ranking question, we assigned each technique an integer score in range [1, 8], giving its subject-ranked position, with 8 meaning "most" and 1 meaning "least". Figures 10-12 and Table 2 summarize results from all subjects. We again used Friedman and Wilcoxon signed-rank tests (as in Section 5.1).

Overall preference: Plots suggest overall preferences tended to follow visual appearance more than behavior rankings. Visual technique impacted overall preference, $\chi^2(7) = 86.9$, $p < .001$. Follow-up tests grouped techniques as listed on the next page:

Table 1: Visual techniques ordered by median performance, per dependent variable, from best (top) to worst (bottom). Per technique, p-values are results of comparisons to IH and OH. ** and * indicate statistically-significant differences (p < .05) and trends (.05 <= p < .10), respectively.

	Penetration depth (cm)				Release time (s)				Translation error (cm)		
	Median	vs. OH	vs. IH		Median	vs. OH	vs. IH		Median	vs. OH	vs. IH
IH	1.992	p<.001**		IH	0.411	p=.031**		ST	0.784	p=.032**	p=.629
ST	2.028	p<.001**	p=.491	VB	0.432	p=.009**	p=.125	IH	0.812	p=.005**	
2H	2.150	p<.001**	p=.192	2H	0.446	p=.002**	p=.781	FC	0.826	p=.054*	p=.147
OC	2.276	p=.004**	p=.019**	OC	0.450	p=.057*	p=.099*	2H	0.845	p=.002**	p=.517
VB	2.323	p=.013**	p=.010**	ST	0.476	p=.045**	p=.082*	VB	0.872	p=.382	p=.020**
AR	2.349	p=.008**	p=.004**	AR	0.507	p=.032**	p=.644	AR	0.927	p=.861	p=.006**
FC	2.444	p=.079*	p<.001**	FC	0.512	p=.120	p=.094*	OC	0.982	p=.098*	p=.060*
OH	2.868		p<.001**	OH	0.561		p=.031**	OH	1.012		p=.005**

1. OC and FC: Significantly preferred over both OH and IH.
2. OH, 2H, and AR: no detected difference or trends compared to OH, but significantly preferred over IH.
3. ST and IH: Significantly worse than OH, but no significance or trends compared to IH (noting borderline ST result).
4. VB: Significantly worse than both OH and IH.

Visual appearance: Visual technique impacted appearance rankings, $\chi^2(7) = 79.4$, $p < .001$. Comparing techniques to IH and OH groups them similarly to the overall preference, except: FC classification hinges on a trend, and ST belongs in the same group as OH, 2H, and AR (preferred over IH, matching our expectation).

Behavior effect: Visual technique impacted behavior rankings, $\chi^2(7) = 51.1$, $p < .001$. Subjects believed OH affects their behavior less than every other technique. This is consistent with penetration results from the targeted ball-drop experiment showing that subjects used lighter touch in other techniques. Compared to IH, subjects believed VB and 2H affect them more. This matches our expectation about strong behavior effect of VB, but VB had more penetration than IH. The VB effect may be more complex than anticipated, as suggested by subject comments: some subjects grasped lighter to avoid intense vibrations, but some others were uncertain about grasp state due to vibration when they were holding the ball, which might discourage light touch.

5.3 Summary Recommendations

5.3.1 Tradeoff between Performance and Preference

OH is a standard approach to visually mimic real-world grasping, but it appears to be the worst performer. IH ranks among the best performers, but is subjectively bad. The goal of additional visual feedback is to balance this tradeoff.

5.3.2 Most Promising Alternatives to OH

We suggest that grasp techniques use 2H when they favor performance and use OC or FC when they favor user experience (or possibly, when combined with other mechanisms that improve performance). Each of them offers notable improvement over OH.

2H has consistently promising performance. It ranks with OH in terms of overall subjective preference (above IH). The assessment does not hinge on any borderline or mixed results. Subject comments suggested a majority of subjects liked 2H, but some subjects disliked its appearance. Future work could study more ways of rendering 2H to seek subjective improvements, but the design study already chose from some alternatives. We consider 2H the best overall technique.

OC is subjectively strong (significantly higher visual and overall rankings than OH and IH). Its performance is a compromise between OH and IH (significant for penetration effects, and trends in other performance variables). Future work could investigate more color variations to seek performance improvements. We already used best tuned scale and interpolation type based on asking subjects to judge effects on grasp control. Objective optimization may find better performance, but could also compromise subjective quality. Combining OC with a release heuristic [7] or haptics, in cases where grasping systems allow this, may reduce any relative weakness of OC.

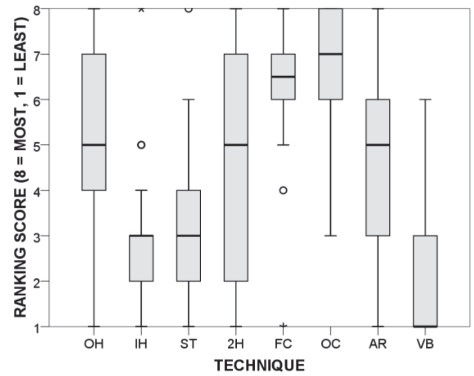

Figure 10: Ranking scores for the “overall preference” question (box-and-whiskers plots; circles and asterisks denote outliers).

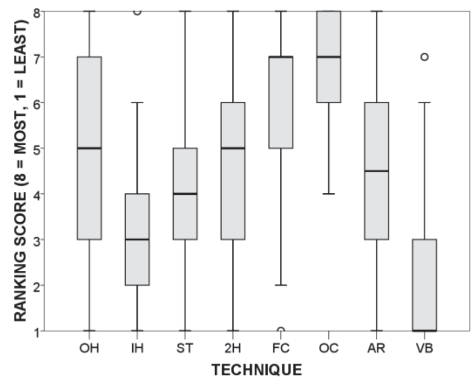

Figure 11: Scores for “how much you liked their appearances”.

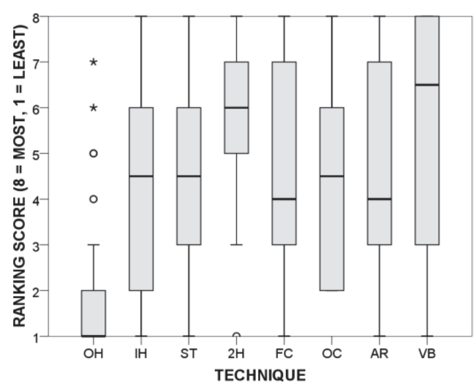

Figure 12: Scores for “how strongly they affected your behavior”.

Table 2: Ranking of techniques from highest (top) to lowest (bottom) median scores, for each question. Per technique, p-values are results of pairwise comparisons to OH and IH. ** and * indicate statistical significance (p < .05) and trends (.05 <= p < .10), respectively.

Overall preference				Visual appearance				Behavior effect			
	Median	vs. OH	vs. IH		Median	vs. OH	vs. IH		Median	vs. OH	vs. IH
OC	7	p=.004**	p<.001**	OC	7	p=.001**	p<.001**	VB	6.5	p<.001**	p=.055*
FC	6.5	p=.023**	p<.001**	FC	7	p=.070*	p<.001**	2H	6	p<.001**	p=.005**
OH	5		p=.001**	OH	5		p=.004**	ST	4.5	p<.001**	p=.193
2H	5	p=.885	p=.006**	2H	5	p=.975	p=.002**	OC	4.5	p<.001**	p=.934
AR	5	p=.501	p=.003**	AR	4.5	p=.455	p=.007**	IH	4.5	p<.001**	
ST	3	p=.022**	p=.100	ST	4	p=.150	p=.028**	FC	4	p<.001**	p=.548
IH	3	p=.001**		IH	3	p=.004**		AR	4	p<.001**	p=.501
VB	1	p<.001**	p=.031**	VB	1	p<.001**	p=.028**	OH	1		p<.001**

FC gave overall similar results to OC, but with weaker support for lighter touch and visual appearance compared to OH (hinging on trends). We additionally compared FC to OC, with no additional findings (release time, $p = .393$; translation, $p = .861$; penetration, $p = .106$; behavior, $p = .206$; appearance, $p = .118$; preference, $p = .374$). Thus, aside from higher confidence in OC results, we do not recommend one over the other.

5.3.3 Remaining Visual Techniques

AR, like OC, provides a performance improvement over OH but appears worse than IH. Although AR subjectively ranked with OH, OC offers a more promising alternative subjectively.

ST may be worthy of further consideration due to good performance, but it is less preferred than 2H, which also has good performance. Performance results of ST, 2H, and IH suggest that future design of additional techniques should consider different ways of revealing the inner (real) hand state directly.

VB had no advantage over IH. We speculated that VB could have unpleasant visuals but strongly encourage light touch. Subjective ranking of behavior effects shows that subjects also expected large behavior effects. However, VB performance gave mixed results, with worse penetration than IH, and VB is subjectively worse than IH.

5.3.4 Other Implications

Results suggest the standard guidelines about avoiding penetrating visuals, e.g., [15], and supporting studies, e.g., [13, 14], should be refined to understand limits of applicability. Multiple techniques that reveal penetration are good for grasp performance, and there may be other interaction types with similar results, e.g., [18].

Work comparing haptic to visual cues for grasping, e.g., [10], should also be re-assessed if it did not use the best-performing visuals. Relative performance of techniques from our study provides a starting point to estimate if results would still hold.

ST results may indicate what can be expected from augmented reality grasping visuals that show real hand state "through" virtual objects, e.g., [3], although the visuals are not exactly identical.

6 Conclusion and Future Work

We tuned and evaluated eight visual feedback techniques for grasping. Compared to just showing a constrained hand, we can improve performance or subjective experience. The techniques giving best performance are not the same as those giving best subjective results, but reasonable tradeoffs can be found. Future work can consider combination with other techniques such as haptic feedback or release heuristics.

References

[1] A. Gomes de Sá and G. Zachmann. Virtual Reality as a Tool for Verification of Assembly and Maintenance Processes. *Computers & Graphics*, 23(3):389-403, 1999.

[2] M. Moehring and B. Froehlich. Effective Manipulation of Virtual Objects Within Arm's Reach. *Proc. of IEEE VR*, pp. 131-138, 2011.

[3] O. Hilliges, D. Kim, S. Izadi, M. Weiss, and A. D. Wilson. Holodesk: Direct 3D Interactions with a Situated See-Through Display. *Proc. of ACM CHI*, pp. 2421-2430, 2012.

[4] L. Kohli. Warping Virtual Space for Low-Cost Haptic Feedback. *Proc. of ACM I3D*, pp. 195, 2013.

[5] D. Kim, O. Hilliges, S. Izadi, A. Butler, J. Chen, I. Oikonomidis, and P. Olivier. Digits: Freehand 3D Interactions Anywhere Using a Wrist-Worn Gloveless Sensor. *Proc. of ACM UIST*, pp. 167-176, 2012.

[6] A. D. Wilson, S. Izadi, O. Hilliges, A. Garcia-Mendoza, and D. Kirk. Bringing Physics to the Surface. *Proc. of ACM UIST*, pp. 67-76, 2008.

[7] M. Prachyabrued and C. W. Borst. Virtual Grasp Release Method and Evaluation. *International Journal of Human-Computer Studies*, 70(11):828-848, 2012.

[8] C. W. Borst and A. P. Indugula. A Spring Model for Whole-Hand Virtual Grasping. *Presence: Teleoperators and Virtual Environments,* 15(1):47-61, 2006.

[9] M. Prachyabrued and C. W. Borst. Visual Interpenetration Tradeoffs in Whole-Hand Virtual Grasping. *Proc. of IEEE 3DUI*, pp. 39-42, 2012.

[10] L. Fabiani, G. Burdea, N. Langrana, and D. Gomez. Human Interface Using the Rutgers Master II Force Feedback Interface. *Proc. of IEEE VRAIS*, pp. 54-59, 1996.

[11] C. W. Borst and R. A. Volz. Evaluation of a Haptic Mixed Reality System for Interactions with a Virtual Control Panel. *Presence: Teleoperators and Virtual Environments*, 14(6):677-696, 2005.

[12] M. Prachyabrued and C. W. Borst. Effects and Optimization of Visual-Proprioceptive Discrepancy Reduction for Virtual Grasping. *Proc. of IEEE 3DUI*, pp. 11-14, 2013.

[13] R. W. Lindeman, J. L Sibert, and J. N. Templeman. The Effect of 3D Widget Representation and Simulated Surface Constraints on Interaction in Virtual Environments. *Proc. of IEEE VR*, pp. 141-148, 2001.

[14] E. Burns, S. Razzaque, A. T. Panter, M. C. Whitton, M. R. McCallus, and F. P. Brooks Jr. The Hand Is More Easily Fooled than The Eye: Users Are More Sensitive to Visual Interpenetration than to Visual-Proprioceptive Discrepancy. *Presence: Teleoperators and Virtual Environments*, 15(1):1-15, 2006.

[15] W. Stuerzlinger and C. A. Wingrave. The Value of Constraints for 3D User Interfaces. *Virtual Realities*, Springer, pp. 203-223, 2011.

[16] R. Boulic, S. Rezzonico, and D. Thalmann. Multi-Finger Manipulation of Virtual Objects. *Proc. ACM VRST*, pp. 67-74, 1996.

[17] G. Zachmann and A. Rettig. Natural and Robust Interaction in Virtual Assembly Simulation. *Proc. of the 8th ISPE Int. Conf. on Concurrent Engineering: Research and Applications*, 2001.

[18] P. Durlach, J. Fowlkes, and C. Metevier. Effect of Variations in Sensory Feedback on Performance in a Virtual Reaching Task. *Presence: Teleoperators and Virtual Environments*, 14(4):450-462, 2005.

[19] T. Ullmann and J. Sauer. Intuitive Virtual Grasping for Non Haptic Environments. *Proc. of IEEE Eight Pacific Conf. on Computer Graphics and Applications*, pp. 373-457, 2000.

[20] Z. Rusak, C. Antonya, I. Horvath, and D. Talaba. The Role of Visual Feedback in Interactive Grasping Simulation. *Proc. of Int. Conf. on Engineering Design*, pp. 6-347-6-358, 2009.

[21] A. Pusch, O. Martin, and S. Coquillart. Effects of Hand Feedback Fidelity on Near Space Pointing Performance and User Acceptance. *Proc. of IEEE Virtual Reality Innovation*, pp. 97-102, 2011.

[22] B. Lok, S. Naik, M. Whitton, and F. P. Brooks, Jr. Effects of Handling Real Objects and Self-Avatar Fidelity on Cognitive Task Performance and Sense of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 12(6):615-628, 2003.

[23] J. Sreng, A. Lécuyer, C. Mégard, and C. Andriot. Using Visual Cues of Contact to Improve Interactive Manipulation of Virtual Objects in Industrial Assembly/Maintenance Simulations. *IEEE Trans. on Visualization and Computer Graphics*, 12(5):1013-1020, 2006.

[24] J. Jacobs, M. Stengel, and B. Froehlich. A Generalized God-Object Method for Plausible Finger-Based Interactions in Virtual Environments. *Proc. of IEEE 3DUI*, pp. 43-51, 2012.

[25] D. Holz, S. Ullrich, M. Wolter, and T. Kuhlen. Multi-Contact Grasp Interaction for Virtual Environments. *Journal of Virtual Reality and Broadcasting*, 5(7), 2008.

[26] Virtual Technologies, Inc. *CyberGlove™ User's Manual*, Sep, 1994.

[27] C. W. Borst and M. Prachyabrued. Nonuniform and Adaptive Coupling Stiffness for Virtual Grasping. *Proc. of IEEE VR*, pp. 35-38, 2013.