Perception of Physical Properties with Augmented Auditory-Tactile Feedback

Master Thesis

GEERT LUGTENBERG
ICA-5576156

MSc Game and Media Technology
UTRECHT UNIVERSITY

Supervisors:	Dr. Wolfgang Hürst
	Assoc. Prof. Christian Sandor
Second examiner:	Prof. Remco Veltkamp

23 MARCH 2016

Abstract

By means of Augmented Reality systems it has become increasingly easy to add virtual objects to our reality and interact with them. But AR can also be utilized to change real objects, i.e. our perception of real objects. In this thesis we present two methods for changing how we perceive thickness when we are exploring objects by touch; by modulating auditory- and tactile feedback.

Vision being a primary sense of humans, we often first estimate physical properties of objects and materials around us by sight. However, when we cannot get the information we want merely by sight, we often resort to haptic exploration. In this research we look at one of such haptic explorations, namely tapping an object to estimate the thickness of a material. We proved that perception of thickness of stiff objects can be changed by modulating sound stimuli. For more flexible objects where tactile cues are more pronounced, adding low frequency long decay tactile stimuli to vibratory responses of tapping on thick objects can make people perceive it as thin. Furthermore we identified that in this last case congruent sound stimuli do not enhance the illusion.

Table of Contents

1. Scientific Paper

This section contains an unpublished scientific paper that is the main deliverable of this thesis. Following it will be a short annotated appendix that goes deeper into some aspects of the research.

Changing Perception of Thickness by Augmenting Auditory-Tactile Feedback

Geert Lugtenberg
Utrecht University
Utrecht, The Netherlands

ABSTRACT

Vision being a primary sense of humans, we often first estimate physical properties of objects and materials around us by sight. However, when we cannot get the information we want merely by sight, we often resort to haptic exploration. In this research one of such haptic explorations is studied, namely tapping an object to estimate the thickness of a material. Results indicate that perception of thickness of stiff objects can be changed by modulating sound stimuli. For more flexible objects where tactile cues are more pronounced, adding tactile stimuli to vibratory responses of tapping on thick objects can make people perceive it as thin. Furthermore we identified that in this last case congruent sound stimuli do not enhance the illusion.

CCS Concepts

•**Human-centered computing → Mixed / augmented reality;**

Keywords

augmented reality; multimodality; physical property perception; haptic interaction

1. INTRODUCTION

It has become more and more common to utilize Augmented Reality (AR) in order to add virtual objects to our environment. Hardware for graphics rendering on a (see-through) head worn display keeps getting better so that computer-generated imagery (CGI) projected onto the real environment go towards uncanny realism. Comparable research in, and hardware for, other modalities is relatively lacking. Though our perception is based on integration of multiple senses, e.g. auditory-tactile [3, 1, 10, 4], sound [16], smell and taste [13]. Exploration of new objects and materials by sight is fast but not all available information is gained this way, or the information can be misleading. An apparently sturdy wooden table can, upon touching it, be a table of thin synthetic material that just has the visual texture of wood.

By picking something up we feel its weight and can make an estimation of the material and the volume. Haptic exploration of real objects, such as picking them up, scratching a surface or tapping them, results in vibrations in the object's material. These vibrations are dependent on certain object and material properties, some of which we estimate when we feel or hear vibrations. We know from elasticity theory [9] that the displacement of a rod or plate that force acts upon depends on the force, elastic modulus of the material and the thickness. There has been previous work investigating stiffness perception (inverse of elasticity) [3, 17] for virtual objects. In this research we expand on this by looking at real objects in an AR setting and the influence of sound and touch feedback upon haptic exploration of *thickness*. Our aim is to:

Change perception of thickness when tapping by modulating auditory-tactile feedback.

We first identify that lower frequencies, high amplitude and long decay of vibrations is characteristic to thinner plates, as opposed to high frequencies with low amplitude and short decay for thick plates. In a first exploratory experiment we find that for stiff objects correct identification of thickness is significantly lower when auditory cues are removed then when tactile cues are removed. This effect is less significant with more flexible material.

In section 3 we show that perception of hollowness (a property resulting from thickness) can be changed by merely generating sounds upon tapping a hollow or solid cube. In section 4 we show that thickness perception can also be altered for objects that have a vibratory response noticeable by touch. Here we look at the influence of both augmented sounds and augmented tactile stimuli when tapping wooden plates of different thicknesses. The results of this experiment indicate that tactile vibratory response has a significant influence on thickness perception. Furthermore we show that augmenting high intensity vibrations with a long decay upon tapping of a plate consistently makes participants perceive it thinner than it actually is.

In addition to these findings we propose auditory-tactile software and hardware as a framework to change the vibratory response to tapping of real objects on the fly. We discuss future work that is potentially interesting using this framework.

2. RELATED WORK

2.1 Perception with haptic interaction

There has been previous work wherein perception of physical properties related to haptic interaction are influenced. These works are often in Virtual Reality and strife to mimic a real sensation in the virtual environment [15, 3]. These sensations can be 'built from the ground-up' whereas real objects have inherit physical properties that are often perceived in a multimodal way. Hachisu et al. [5] were able to add or subtract vibrations to a haptic pen when it applies force to a material. This augmentation creates the illusion of touching a different material than the real one. Inspired by this work we propose a similar method for augmenting vibro-tactile feedback in the second experiment described in section 4. In a surprisingly easy setup, TECHTILE toolkit [11] can play tactile sensations realtime or recorded for a very robust haptic illusion, e.g. balls rolling around in a cup. We will utilize this also in the second experiment.

2.2 Multisensory integration

It is possible to get a desired effect of perception of touch by stimulation of another modality. This influence of a seemingly unrelated modality is called multisensory integration. A famous and robust illusion of weight is the 'Size-Weight illusion' [2]. Visual indication of size has a cross-modal effect on the perception of weight of two equally heavy objects. When stroking a surface 'roughness' can be changed with visual [8, 7] and auditory [4] stimulation. Hardness perception can be changed by deforming CGI of a texture that is projected on a surface upon pressing it [14, 6]. We are interested in whether sound and tactile stimuli also have an integration effect on thickness perception.

3. EXPERIMENT 1: AUGMENTED SOUND

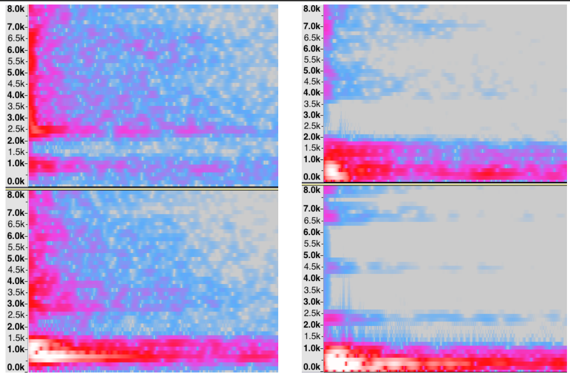
We divided experiments into two: in the first we get a general sense of the relative influences of hearing and touch when tapping an object to determine thickness. Then we augment the auditory modality and measure correctness of discriminating between thicknesses. In the second experiment (section 4) we augment the tactile modality, and finally we augment both auditory and tactile to see if there is a multimodal integration effect on thickness perception.

We chose to investigate augmenting sound before tactile vibrations, because auditory hardware is simpler and easier to obtain, i.e. headphones can be found anywhere and everywhere but haptic hardware is much more limited. We assume that if there is no tactile indication of thickness, i.e. no noticeable difference when touching a thick and thin object, perception of thickness is entirely dependent on what we hear. Therefore we can change perception of thickness solely by modulating sound feedback. We summarize this assumption in our first research question:

RQ 1 *Can we achieve a different perception of thickness (solid or hollow) by solely modifying auditory feedback during tapping, but keeping the object's physical properties fixed?*

We approach this problem by measuring correct identification of 'solid' or 'hollow' cube in a yes/no type psychophysical experiment.

We compared the sound spectrograms of impulse responses of thick and thin materials. As can be seen in Figure 1, a long decay of low frequencies is characteristic for thin material, as opposed to short decay of higher frequencies. We generate sound by resynthesizing the original sound of a tap on a cube in real-time and transforming it according to the desired characteristic (Table 1). By resynthesizing we take into account the acoustical properties of the material and object (modal models [16]), and the velocity and duration of the tap.



(a) (b)

Figure 1: Spectrograms of impulse responses of (a) plastic cubes and (b) wooden plates show characteristics of thick (top row) and thin (bottom row) material. Sounds originating from thin objects have a high amplitude, shown in white, in the lower frequency range. Furthermore it is shown that these high amplitude low frequencies have a longer decay time.

Table 1: Characteristics of thickness in vibrating material

	Avg amp	Frequency range of high amps	Decay time
Thick	Low	1000 Hz ~	short
Thin	High	0 Hz ~1000 Hz	long

3.1 Platform

A Mogees piezoelectric sensor [12] is connected with the Mogees Virtual Studio Technology (VST) plugin running in Cycling '74 Max 6 on a MacBook Pro. The platform can be seen in Figure 2. The VST plugin transforms the vibrations measured by the piezoelectric sensor in real time and outputs it as audio. We use two 3D printed cubes of a styrenic plastic material that have a Young's Modulus of 2.0-2.6 GPa. The cubes' dimensions are 70x70x70mm and the hollow cube is 2mm thick. Sehnheiser CX3.00 in-ear headphones are used to play audio, so that earmuffs can be worn over the headphones to cancel external sound.

3.2 Method and Procedure

Before participants enter the examination room, a cube of variable thickness (solid or hollow) is placed in the holder. Participants read an info sheet after which an examiner explains the procedure again verbally. Each participant tests three conditions: no sound augmentation (control), white noise sound (no auditory cues) and resynthesized sound (thin or thick).

The resynthesized sounds are pitched down and up, and

Figure 2: Setup for experiment 1: (left) Participants tap the top of a solid or hollow cube. There is visually no distinction between the two thicknesses (the opening in the right figure is faced down). The cube is placed in a holder that is firmly attached to a table surface. A piezoelectric sensor is attached to the side of the cube (right) to capture vibrations in the material.

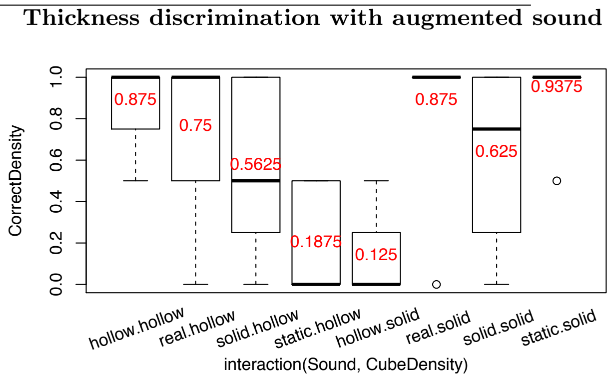


Figure 3: Boxplot of results of the thickness discrimination task in experiment 1. Mean scores of correctly identifying the cube's real thickness are indicated in red.

decay time is increased and decreased for thin and thick modulation respectively. The values for pitch and decay are determined subjectively in order to create a big contrast between thin and thick sound.

Participants are asked to only tap the cube in front of them on the top side, with the index finger of their dominant hand. They are asked to decide whether the cube is hollow or solid, and until they do they can tap as many times as desired. After a participant decides this, the cube is changed for a different thickness. The participants write their decision on an answer sheet that is behind a wall; exchange of the cube thickness is not seen. Cube thickness and resynthesized sound where counterbalanced over participants.

In summary, independent variables are *cube thickness* (solid, hollow) and *tapping sound* (real, static noise, thin, thick). Each participant repeats all permutations of these variables one time, resulting in 16 trials per participant. The duration per participant was approximately 30 minutes.

3.3 Participants

This first experiment is meant to identify trends in perception of thickness. We do not require statistical significance but are interested in a direction to follow-up our research question, i.e. if sound is sufficient (RQ1) or tactile augmentation is necessary. Therefore, we base the number of participants on the number of unique permutations of the independent variables: 8 participants. Participants are students with ages ranging from 20-31.

3.4 Results and discussion

The results of this experiment are visualized in Figure 3. We observe that in the control condition, where real sound is heard, participants scored on average 81% correct. When any sound cue of thickness is removed by generating static noise, correctness for solid cubes stay high (94%) whereas correctness for hollow cubes drops to 19%. Similarly with hollow sound stimuli, correctness for hollow cubes is high (88%) and for solid cubes is low (13%). Correct discrimination of thickness when sound is augmented solid is by chance: 56% and 63% for hollow and solid cubes respectively.

We observe that participants consistently answer 'hollow' when presented with hollow sound cues (88%). The same is not true for solid sound cues with 53%. It is possible that 'solid' sound that was generated is not indicative enough by itself. If participants would compare the 'solid' sound to 'hollow' sound, it is likely they will discriminate between the two correctly. In experiment 2 we account for this by changing the experimental design to a two-alternative forced choice task. Furthermore participants reported that the augmented sound did not match any material they knew, and thus found thickness discrimination very difficult. We see in the results of thickness discrimination with real sound that correctness is high. We speculate that improving the sound quality and 'recognizability' of material will result in a higher percentage of answers that are equal to the sound heard for both hollow and solid.

It is clear from the results that perception of thickness of two objects that have no noticeable tactile difference when you tap on them, can be changed by modulating response sound. Specifically, when forced to classify an object as 'hollow' (thin) or 'solid' (thick), sounds that are indicative of thin material make participants classify it as 'hollow' consistently.

4. EXPERIMENT 2: AUDITORY-TACTILE

Experiment 1 showed that sound can change thickness perception, when there is no indicative tactile sensation noticeable. In this experiment we changed the cubes for larger wooden plates in order to make this tactile indication of thickness more pronounced. Our aim is to measure the influence these two modalities have in thickness perception. Furthermore we augment both sound and touch feedback when participants tap to measure if there is an integration between the two senses. This is summarized in the second research question:

RQ 2 *When there are both auditory and tactile indications of thickness, can we change perception of thickness by modulating auditory-tactile feedback during tapping?*

Because 'thin' and 'thick' are context specific terms and subjective we do not repeat the experimental design of experiment 1. Instead we measure correct identification of the 'thinner' of two plates in a two-alternative forced-choice (2AFC) psychophysical experiment.

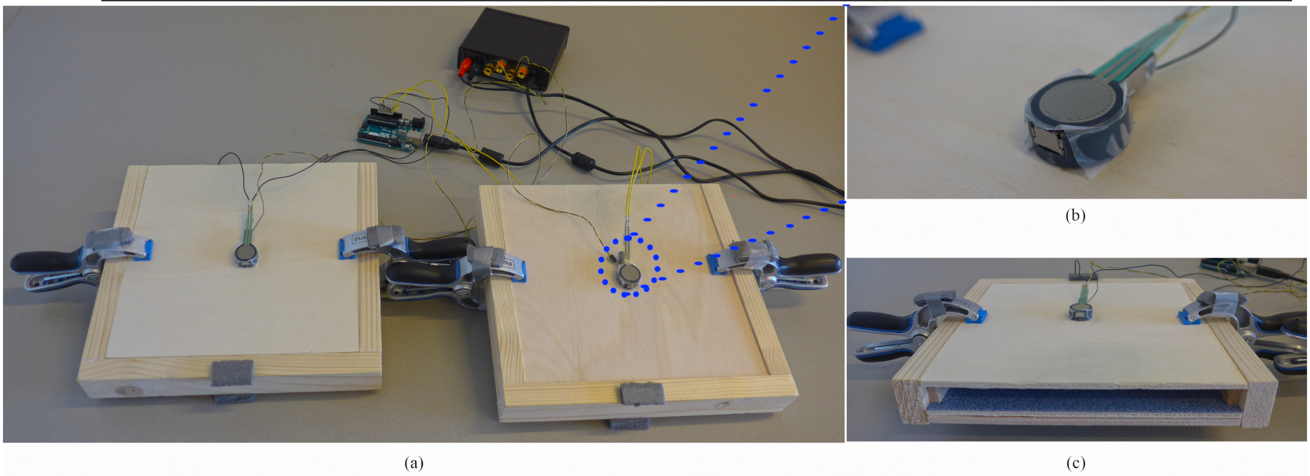


Figure 4: The setup for experiment 2: (a) two wooden plates are clamped into two wooden frames. Vibration-damping material is attached to places where the wood or clamps touch the underlying surface. In the middle of the plates are Force Sensing Resistors (FSR) that are pasted on top of vibro-tactile actuators, as can be seen in more detail in (b). The actuators are connected to an amplifier and together with the FSRs make a serial connection to a computer. (c) is the back-view of the frames: Both frames can be turned-over to change the thickness of the plates.

4.1 Platform

Two wooden frames were constructed to each hold two wooden plates in place. The wooden plates are of dimensions 200x200xTmm where T=4mm (thin) and T=10mm (thick). Each frame can hold a thick and a thin plate, and has two clamps in order to clamp them in place. This prevents inter-resonance. See Figure 4 (a) and (c). Both the clamps and the wooden frames are covered with 9mm thick Polyethyleen foam, to stop vibrations in the plates from resonating into the underlying surface. To generate tactile vibrations upon touching the plates we created two tactile sensor/actuators (Figure 4 b). These consist of a PVC cylinder with a diameter of 20mm and height 8mm with a cutout in which our vibro-actuators fit precisely. The vibro-actuators and connected AL-202H Amulech amplifier are per design of TECHTILE toolkit [11] and can display a range of 1-20000Hz. A Force Sensitive Resistor (Interlink FSR402) is attached on top of the vibro-actuator and PVC cylinder. Pressure on the FSR is registered by our serial connected software and measures a tap force approximately between 0 and 50N. Based on the tap force our software plays a prerecorded audio file. Measured latency between moment of impact on the FSR and audio output ranges from 20ms to 60ms.

4.2 Method and Procedure

In this experiment each participant is tested under three conditions of augmented modalities: tactile stimuli (thin, thick), auditory-tactile stimuli (congruent with each other), auditory-tactile stimuli (discrepant with each other). Under each condition real tactile feedback of the plates is present. The participants are asked to determine the *thinner* of two plates before them by tapping on the sensors. The two plates are always of opposite thickness; one thick and one thin. However the location of the thin plate is randomized each round, resulting in 50% left and 50% right. There are no visual cues that indicate thickness.

In section 3.4 we discussed that participants did not recognize a material from sound cues generated by our system. To make sounds optimally recognizable as belonging to a material we capture tapping sounds upon the real material in a preprocessing phase. While capturing the sound we also capture the impact force of the tap and save it as a pair to be played-back during the experiment. We captured 12 sound-force pairs of tapping on thick wood, and 9 sound-force pairs of tapping on thin wood. The forces of the taps are approximately equally distributed ranging from subjectively 'soft touch' to 'very hard tap'.

We chose to generate vibro-tactile feedback by modulating a sine wave with specific characteristics as in Table 1. Our aim is not to simulate all touch sensations, as this will arguably result in a virtual reality problem. Instead we add vibratory cues to vibrations in the real material. Frequencies, base amplitudes and decay times for thick and thin vibrations are summarized in Table 2. Decay time means the total time for an exponential fade-out to zero amplitude of the sine wave. These specific values were chosen subjectively to have a large contrast between 'thick' and 'thin'.

Upon entering the examination room, participants were asked to read an information sheet and not to touch anything on the setup. Every participant went through one training round to familiarize themselves with the auditory and tactile experience. Also tapping intensity and order was practiced. Participants had to tap the left plate three times on the sensor: soft tap (place finger upon the sensor, $\pm$ 1N force), medium tap ($\pm$ 20N force), hard tap ($\pm$ 50N force). For every tap they had to use the index finger of their dominant hand and start with the left plate, then the right plate. After this process participants indicate which plate they think is thinner by saying 'left' or 'right'. They then turn around in their chair, facing a back wall, so that they do not see the examiner change the permutation of thickness of the plates. After the examiner logged the answer and changed the thickness permutation, the participant was tapped on the shoulder and the process is repeated in a new round. During the training round participants tapped

Table 2: Tactile vibratory properties

	Base amp	Base freq	Decay time
Thick	0.3	500Hz	100ms
Thin	1.0	100Hz	500ms

on the plates without any augmentation, and asked which plate was thinner. They were then told if they were correct. While wearing the in-ear headphones they were presented with a sample of thick and a sample of thin tapping sound on wood. They had to indicate which of the two sounds 'sounded thinner'. After this training round, participants are asked to wear earmuffs over the in-ear headphones, after which normal trials started. Every permutation of the independent variables was repeated ten times, resulting in 60 trials per participant. The total duration per participant was approximately 45 minutes.

4.3 Participants

Ten male students participated in this experiment, with ages ranging from 20 to 26. One participant had nails sufficiently long that it touched the sensor when tapping. In this specific case the participant was asked to tilt his finger such that the nail would not touch the sensor. This variation did not show up as outliers in the results.

4.4 Results and Discussion

Figure 5 visualizes the mean correctness scores for determining the thinner of two plates. The x-axis labels indicate a condition where only tactile cues where augmented (H.0, H.1), tactile and sound cues were augmented and thickness of the cues were congruent (C.0, C.1) or discrepant (D.H, D.S). Furthermore the number after the modality labels indicate whether thickness of the stimuli match the real thickness (1) or not (0). 'D.S' means a discrepancy between tactile and sound thickness stimuli, where the sound thickness matches the real thickness. For 'D.H' tactile thickness matches the real thickness.

The results are analyzed with a two-way ANOVA with repeated measures and an error term on differences between participants. Normality of the mean data was confirmed using a Shapiro-Wilk test and QQ-normals plot. The results of the ANOVA analysis are: there is a significant effect of sound on correct discrimination of thickness ($p = 0.0959$) and a significant effect of tactile feedback ($p = 0.0441$). However an interaction effect between the two was not significant ($p = 0.492$).

We aimed to answer RQ2: is it possible to change perception of thickness when auditory and tactile cues of thickness are noticeable. Observing the results in Figure 5 we see that discrepant auditory and tactile cues lower correct identification of a thin plate to randomly chosen by chance. However, when there is only tactile feedback (H.1 and H.0) or congruent auditory-tactile feedback (C.1 and C.0) perception of thickness can be somewhat 'controlled', i.e. when presented with tactile 'thin' stimuli upon tapping a thick plate and tactile 'thick' stimuli upon tapping a thin plate, 76% of answers (1 - mean of H.0) of participants perceive the thick plate to be thinner than the thin plate.

5. CONCLUSION AND FUTURE WORK

In this paper we aimed to get a better understanding of the influences that hearing and touch have when tapping an object. The motivation for this is that it is intuitive to touch objects to get information about material and thickness, when this is not clear by looking at them. Furthermore we investigated if it is possible to augment auditory feedback, tactile feedback or a combination of these modalities in order to change perception of thickness. In two experiments we have shown that perception of thickness can be altered by modulating auditory and tactile feedback. Stiff thick objects can be perceived to be thin by only augmenting thin sound cues upon tapping. In the case of more flexible objects where tactile cues are more pronounced, perception of thickness can be changed by only modulating tactile feedback. By adding tactile low frequency, long decay cues to thick objects, they are perceived to be thin. Furthermore we found no significant interaction between auditory and tactile feedback in this case.

This study is novel in testing thickness perception with auditory-tactile stimulation and an addition to ongoing work of changing perception of physical properties. Methods and the auditory-tactile feedback system described here can be used in future research into perception of vibrations as a result of haptic interaction. An idea to further this research is to introduce visual cues of material texture and different material sounds and repeat the second experiment described. It is possible that an expectation of material changes the results presented here.

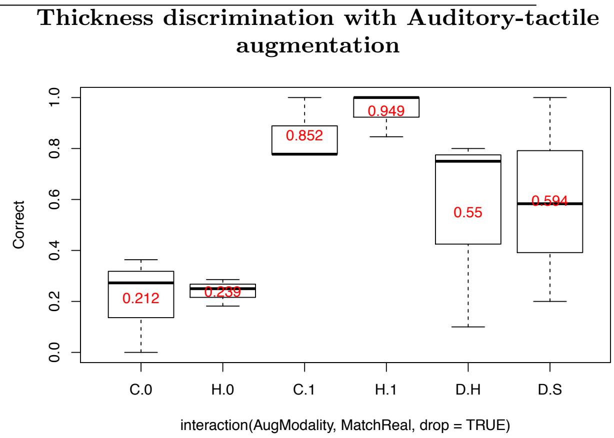


Figure 5: Box plots of experiment 2 results with means indicated in red. The labels C indicate Congruency between tactile and sound cues, labels D for Discrepancy. Labels H indicate only tactile modality. '1' indicates that the thickness simulated by the augmented modalities matches the real plate thickness, '0' indicates a mismatch.

6. REFERENCES

[1] F. Avanzini and P. Crosato. Integrating physically based sound models in a multimodal rendering architecture. *Computer Animation and Virtual Worlds*, 17(3-4):411–419, 2006.

[2] A. Charpentier. Experimental Study of Some Aspects of Weight Perception. *Archives de Physiologie Normales Pathologiques*, 3:122–135, 1981.

[3] D. E. DiFranco, G. L. Beauregard, and M. A. Srinivasan. The Effect of Auditory Cues on the Haptic

Perception of Stiffness in Virtual Environments. In *Proceedings of the ASME Dynamic Systems and Control Division*, volume 61, pages 17–22, 1997.

[4] S. Guest, C. Catmur, D. Lloyd, and C. Spence. Audiotactile interactions in roughness perception. *Experimental Brain Research*, 146(2):161–171, 2002.

[5] T. Hachisu, M. Sato, S. Fukushima, and H. Kajimoto. Augmentation of Material Property by Modulating Vibration Resulting from Tapping. In *Proceedings of the International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 173–180, 2012.

[6] Y. Hirano, A. Kimura, F. Shibata, and H. Tamura. Psychophysical Influence of Mixed-Reality Visual Stimulation on Sense of Hardness. In *Proceedings of IEEE Virtual Reality*, pages 51–54, 2011.

[7] A. Iesaki, A. Somada, A. Kimura, F. Shibata, and H. Tamura. Psychophysical influence on tactual impression by mixed-reality visual stimulation. In *Virtual Reality Conference, 2008. VR'08. IEEE*, pages 265–266. IEEE, 2008.

[8] M. Kagimoto, A. Kimura, F. Shibata, and H. Tamura. Analysis of Tactual Impression by Audio and Visual Stimulation for User Interface Design in Mixed Reality Environment. In *Proceedings of the International Conference on Virtual and Mixed Reality*, pages 326–335, 2009.

[9] L. D. Landau and E. Lifshitz. Theory of elasticity, vol. 7. *Course of Theoretical Physics*, 3:109, 1986.

[10] S. J. Lederman, R. L. Klatzky, T. Morgan, and C. Hamilton. Integrating multimodal information about surface texture via a probe: relative contributions of haptic and touch-produced sound sources. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on*, pages 97–104. IEEE, 2002.

[11] K. Minamizawa, Y. Kakehi, M. Nakatani, S. Mihara, and S. Tachi. Techtile toolkit: a prototyping tool for design and education of haptic media. In *Proceedings of the 2012 Virtual Reality International Conference*, page 26. ACM, 2012.

[12] Mogees Ltd. Mogees pro. http://www.mogees.co.uk/pro, 2016. Accessed: 2016-10-17.

[13] T. Narumi, S. Nishizaka, T. Kajinami, T. Tanikawa, and M. Hirose. Augmented Reality Flavors: Gustatory Display Based on Edible Marker and Cross-modal Interaction. In *Proceedings of Human Factors in Computing Systems*, pages 93–102. ACM, 2011.

[14] P. Punpongsanon, D. Iwai, and K. Sato. SoftAR: Visually Manipulating Haptic Softness Perception in Spatial Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1279–1288, 2015.

[15] N. Rosa, W. Hürst, W. Vos, and P. Werkhoven. The Influence of Visual Cues on Passive Tactile Sensations in a Multimodal Immersive Virtual Environment. In *Proceedings of the ACM International Conference on Multimodal Interaction*, pages 327–334, 2015.

[16] K. Van Den Doel, P. G. Kry, and D. K. Pai. FoleyAutomatic: Physically-based Sound Effects for Interactive Simulation and Animation. In *Proceedings of ACM Computer Graphics and Interactive Techniques*, pages 537–544, 2001.

[17] W.-C. Wu, C. Basdogan, and M. A. Srinivasan. Visual, haptic, and bimodal perception of size and stiffness in virtual environments. *Asme Dyn Syst Control Div Publ Dsc.*, 67:19–26, 1999.

2. Annotated Appendix

In this annotated appendix the distinction between Object Density and Thickness will be explained. Furthermore the process of generating sound for the experiments is expanded upon, as well as why multiple materials were not included in the research.

2.1 On Object Density and Thickness

The first deliverable of this thesis, a position statement paper (section A) in which experiment 1 was proposed, had its focus on perception of object density. This as opposed to the general theme of this thesis, which is perception of thickness. Object density and thickness are arguably two descriptors of volume of an object. We define object density as the mass per unit volume of the *apparent* object, i.e. the visible hull of the object. For example, a hollow object has a lower object density than a solid object. Another definition that could have been used is 'hollowness'. Physical property 'thickness' is used from experiment 2 on however, because it is much more intuitive and has a basis in elasticity theory.

When we touch a plate, as in experiment 2, the information that we get is by feeling and hearing vibrations of the plate. How much a plate vibrates depends on its *flexural rigidity* and force applied:

$$D = \frac{EH^3}{12(1-v^2)}$$

where D is flexural rigidity, H is thickness, E is Young's Modulus and v is Poisson's ratio. This means that from a physics basis we should be able to gather information about the material (Young's modulus and Poisson's ratio) and thickness when we apply force to a plate.

2.2 Methodology for Sound Generation

The method for generating sound in the experiments was not extensively described. Originally three methods for generating sounds were proposed, see Figure 2.1. These methods differed in implementation difficulty, sound realism and robustness/generalization. Initially only the easiest to implement method was used, which was resynthesizing real sound at run-time. It was thought that playing back recorded sounds would result in very poor realism, i.e. tactile sensation would feel separate from auditory sensation. Hardware used was piezoelectric sensor Mogees Pro and accompanying software for resynthesizing, and indeed latency was not noticeable. Mogees Pro software in the form of a VST plugin had certain presets for resynthesizing sound such as damping (decay time) for 4 modal sines and pitch shifting. Subjectively damping and pitch shifting values (Table 2.1) were found for resynthesizing 'hollow' and 'solid' sounds.

Parameter	Hollow sound	Solid sound
Sine 1 Damping	0.80	0.00
Sine 2 Damping	0.80	0.00
Sine 3 Damping	0.00	0.00
Sine 4 Damping	0.08	0.08
Pitch shift	One octave down	One octave up

TABLE 2.1. Parameter values in Mogees Pro re-synthesizer to generate hollow and solid sounds.

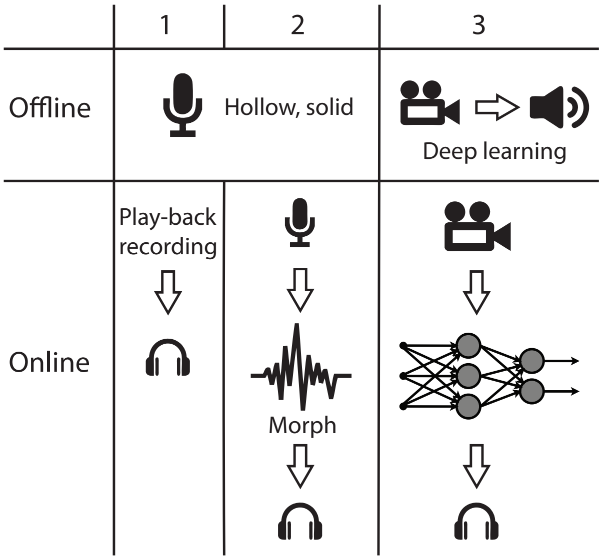


Figure 2.1: Schema of three methods that generate sound on haptic interaction. (1) Pre-record sounds in an offline phase and play them back upon touching the object. (2) Analyze sounds in an offline phase and morph (resynthesize) captured sounds that originate from touching the object to fit the analyzed ones. (3) Deep learning of haptic interaction audio and video to predict sound features with only video input.

Results from experiment 1 indicated that sound quality and recognizability was important in correct thickness discrimination. The third method proposed to generate sound is learning sound features from audio/video input into a deep neural network. This approach is most robust, i.e. given enough training samples, it can predict sound features merely based on video of tapping on different materials. However, we did not expect that generating sounds with the Deep learning approach would increase quality and recognizability and thus we did not pursue this method further.

2.3 Methodology for Experiment 2

The results from experiment 1 indicated that participants were not able to recognize material from the sounds that they heard. We hypothesized at this point that being able to see the colors and texture of a material on which you are tapping might have an influence on thickness perception. Intuitively, if you visually recognize a material as a ceramic you expect it to feel stiff. Furthermore, it creates an expectation of the sound that the material will emit upon tapping it. We wanted to expand experiment 1 to include visuals and sounds of different materials, to study what the influence is of these modalities on thickness perception. We created a similar platform as in experiment 2. Plates had the texture of plastic, wood, marble and aluminum (Figure 2.2). We recorded tapping sounds of different impact force on the aforementioned real materials. Then we executed a small pilot study to see if participants could hear the difference in thickness and feel the difference, and what the influence of different materials was. Two participants repeated all unique permutations of independent variables one time, resulting in 32 observations. In all observations the participants could *feel* the difference between the plates and indicated the thinner plate correctly. Participants reported they did not pay attention to the 'color of the plates' and believed their tactile sense more than their hearing. At this point we decided not to continue with different materials and textures, but with tactile stimuli. It seemed that for more flexible objects, tactile feedback was more distinctive of thickness than sound. Another reason to abandon different material textures was because in our setup it took quite some time to change the material plates. The clamps (as in Figure 2.2) had to be removed, the Force Sensing Resistor removed, the plate changed for another material, the FSP pasted on it again, and the clamps reapplied. During this time the participants had to leave the experimentation room. For all permutations of variables it took more than an hour per participant, where half of the time was spend changing material plates.

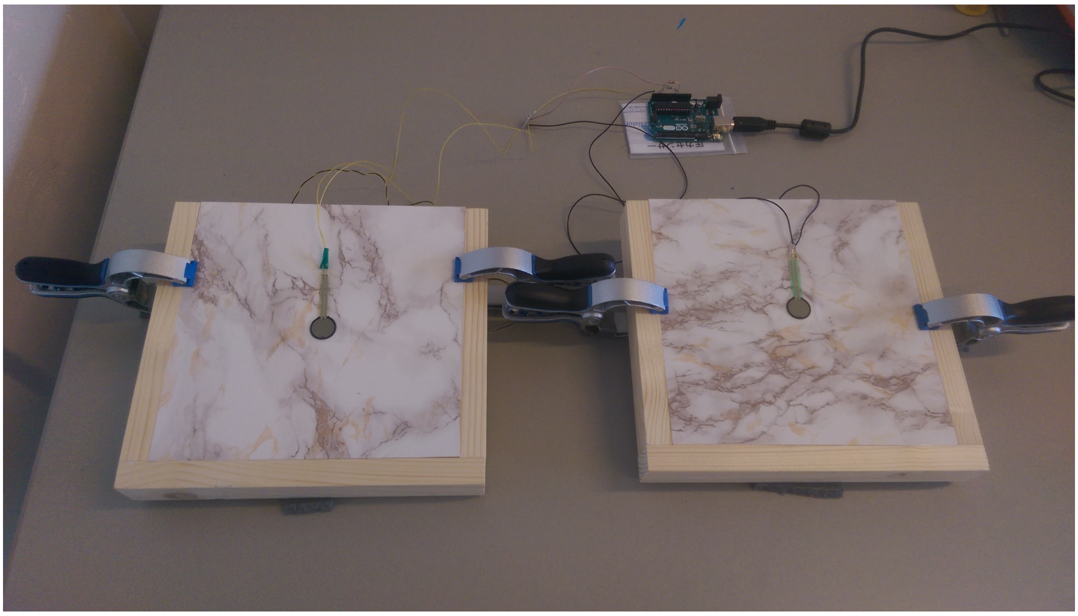


Figure 2.2: Platform for experiment 2 with different materials. Here, two wooden plates have the texture of marble. Upon tapping the sensors in the middle of the plates, pre-recorded sounds of tapping of real marble were heard.

A. MVAR Workshop Paper

This appendix contains a position statement paper that was part of the deliverable for Phase 1. It was published in Proceedings of the 2016 workshop on Multimodal Virtual and Augmented Reality (MVAR) and presented at the workshop during phase 2 of the thesis. MVAR was part of the ACM International Conference on Multimodal Interaction in Tokyo, 2016.

Changing Perception of Physical Properties using Multimodal Augmented Reality: Position Paper

Geert Lugtenberg
Utrecht University
The Netherlands
g.lugtenberg@students.uu.nl

Christian Sandor
Nara Institute of Science and Technology, Japan
sandor@is.naist.jp

Wolfgang Hürst
Utrecht University
The Netherlands
huerst@uu.nl

Alexander Plopski
Nara Institute of Science and Technology, Japan
plopski@is.naist.jp

Takafumi Taketomi
Nara Institute of Science and Technology, Japan
takafumi-t@is.naist.jp

Hirokazu Kato
Nara Institute of Science and Technology, Japan
kato@is.naist.jp

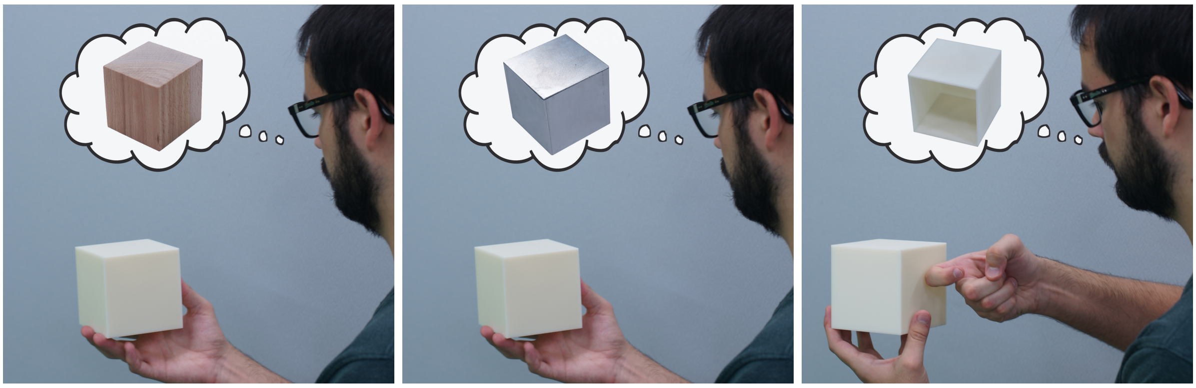


Figure 1: In the ultimate AR system, perception of physical properties, such as material and object density, can be changed as desired.

ABSTRACT

By means of augmented reality (AR) systems it has become increasingly easy to manipulate our perception of real objects. In this position paper we review existing work that changes physical property perception, and propose methods for changing perceived object density during haptic interaction. Our goal is to modulate the sound emitted by an object when touched. We hypothesize that augmented sound can make people think that an object is hollow or solid regardless of its actual object density. We describe an experiment to validate this hypothesis. Participants tap a hollow or solid cube and are asked to determine if it is hollow or solid, based on the multimodal feedback.

This research is a first step towards an AR system that can alter multimodal perception of object physical properties, and open doors for related research.

CCS Concepts

•**Human-centered computing** → **Mixed / augmented reality;**

Keywords

augmented reality; multimodality; physical property perception; object density; audio feedback; haptic interaction

1. INTRODUCTION

A large number of applications in AR to date focus on realistic visual rendering of virtual objects, and changing how real objects are visually perceived. Comparable work in other modalities is relatively lacking. Achieving consistent realism in other modalities, e.g. touch [2], sound [10], smell and taste [7], is still challenging as augmenting a new virtual object into the environment often requires additional hardware (e.g. haptic machinery) and physics simulation. But if an existing object is used, often we can achieve our goal by only manipulating the object's physical properties instead of creating a new object. Figure 1(a,b) shows a visual manipulation of an object's material, and (d) auditory manipulation of an object's density. Depending on the application, only perception of some properties of the object need to be changed, e.g. hardness, color, density, to appear realistic. The ultimate goal of this research is the creation of an AR system that can change perception of object physical properties, in order to simulate desired material and object-density.

With the increased accessibility of 3D printing, studies aimed towards perception of physical properties can contribute to using 3D printed objects as 'templates'. Perception of more and more physical properties of these template objects could be altered as desired. In this position paper we review existing work aimed at changing physical property perception and evaluate what is missing to reach the ultimate goal. Finally, we propose a research plan to study object density perception.

2. RELATED WORK

We review previous work on modifying perception of physical properties of rigid objects. The ultimate AR system could ideally alter perception of all physical characteristics of objects, however characteristics that humans cannot perceive without instruments are of far less interest. We are concerned with perception of physical properties that play a role when haptically interacting with objects. In particular, we focus on stiffness, mass (weight), roughness and hardness.

2.1 Stiffness

Stiffness is the ability of an object to resist deformation in response to force. It is perceived by multiple sensory systems, in particular audition and tacticion. DiFranco et al. [2] explore the influence of sound on the haptic perception of stiffness in a virtual environment. They use a haptic device to simulate feedback from a surface with varying stiffness. Upon contact with the surfaces, impact sounds are generated. They found that audio cues affected the stiffness perception, even when the haptic feedback remained the same. Hachisu et al. [3] created a haptic device that, upon impacting a real object, adds or subtracts vibrations to the resulting surface vibration. They found that this changed the perceived stiffness of the material and created the illusion that the real object was made of a different material.

2.2 Weight

The 'Size-Weight illusion' is a famous illusion of perceived weight [1]. It states that when holding objects of the same weight but different size, the bigger object is perceived to be lighter than expected. R-V Dynamics [4] created an illusion that changes weight perception of a real object. They superimposed computer-generated imagery (CGI) representing inertial force of movable objects in an AR environment. They found that with CGI of moving (dynamic) volumes, objects are perceived to be lighter. Visual stimuli can also affect perceived vibrotactile intensity [9]. Vibrations are perceived 'lighter' when visual cues suggest a larger or heavier object. These works suggest that discrepant perceived weight accompanying a change in perceived density could be countered with visual stimuli.

MVAR'16, November 16 2016, Tokyo, Japan

 ISBN 978-1-4503-4559-0/16/11…$15.00

DOI: http://dx.doi.org/10.1145/3001959.3001968

2.3 Surface roughness

Kagimoto et al. [6] analyze the perceived 'roughness' of a real surface with visual and audio stimuli. They superimpose a material on a surface, and play augmented audio stimuli when the user touches the surface. They found that when two objects have identical roughness, visual and audio stimulation can change perception of haptic roughness between the two. In this research, the roughness of the physical object remains the same and thus visual and auditory cues could be utilized to make the user perceive a rougher or less rough surface.

2.4 Hardness

Perception of hardness is not only a tactile experience but can also be influenced by visual stimulation [5]. In SoftAR [8] the haptic experience of softness is influenced by visual stimulation of the surface that is interacted with. Deformation of the surface of an object is experienced haptically, but also visually. This knowledge is used to project an exaggerated deformed image onto a surface on touch, to simulate increased softness. This technique can be used when simulating material that is softer than the physical material (e.g. perceiving wood or plastic to be styrofoam).

3. PROPOSED RESEARCH

From the previous section we notice a lack in the literature of object density perception. We propose a research plan on changing perceived object density.

3.1 Problem

Density is a material property that represents mass per unit volume. We define *object density* as mass per apparent object volume, i.e. a hollow object has a lower object density than a solid object of the same apparent volume. Object density is perceived by integration of multimodal cues. Applying a force impulse (e.g. tapping the object, scratching the object) causes the object to vibrate. This can be an indicator for density by means of haptic feedback. Less dense objects vibrate more powerfully because there is less mass to displace. These vibrations are perceived not only haptically, but also aurally. Based on this, we hypothesize that augmenting interaction sound feedback can change the perception of object density.

This research focuses on how to manipulate perceived object density. We approach this problem by augmenting the sound heard when tapping on an object (see Figure 2). We further investigate what effect discrepant audio feedback has when discriminating between a hollow and a solid object by asking the following research questions:

RQ 1 *How does auditory feedback influence our multimodal perception of object density?*

RQ 2 *Can we achieve a different perception of object density (solid or hollow) by modifying auditory feedback during tapping, but keeping the object's physical properties fixed?*

3.2 Methodology

We investigate three possible methods to augment real contact sound to simulate object density (Figure 3). In all cases there is an offline phase. The methods are evaluated by their expected performance, realism and implementation difficulty.

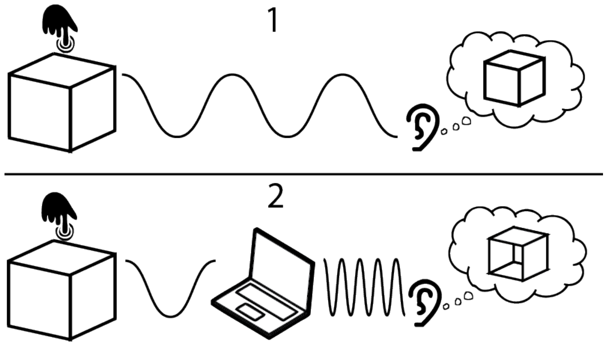


Figure 2: Our approach to modify perceived object density. Haptic interaction with a solid cube produces sound. (1) The cube is perceived as being solid. (2) The sound is modified so that the cube is perceived as hollow.

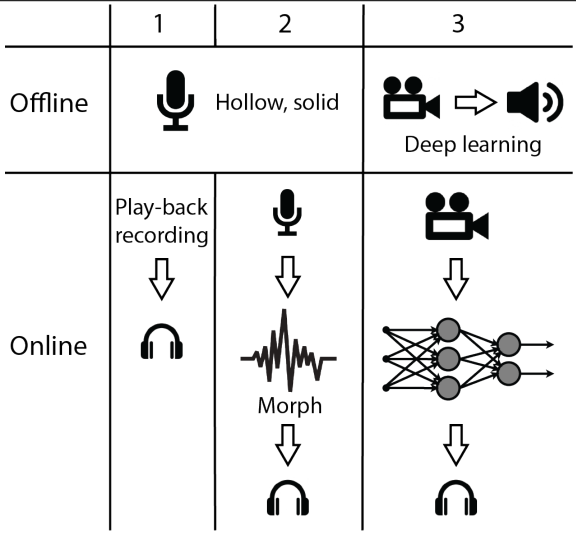


Figure 3: Schema of three methods that augment contact sounds: (1) Pre-record audio, (2) Morph real audio, and (3) Deep learning of interaction A/V. Every method has an offline and an online phase.

3.2.1 Pre-record audio samples

The simplest way to simulate object density with augmented contact audio is to play back the pre-recorded audio samples. In the offline phase, sound emitted by tapping a hollow and solid cube is recorded. At run-time, when the target object is touched the user hears the recorded sound. Sounds are cued by a Wizard of Oz style presentation where the supervisor initiates the playback.

Although this method is easiest to implement, we expect that audio will not have a high level of realism, because of variable force used when tapping and temporal discrepancies between moment of impact and audio sample being played.

3.2.2 Morph captured interaction sound

The second approach is to modify pre-recorded audio samples to captured impact duration and velocity. The preparation phase is the same as for the first method: record tapping and scratching sounds coming from interaction with cubes. A microphone captures vibrations that occur in the object when there is haptic interaction. Based on frequency analysis of the input, sound is morphed to more closely resemble the desired sound.

This method presents a good trade-off between implementation difficulty, run-time performance and realism. We expect it to perform better than playback of pre-recorded audio samples.

3.2.3 Deep learning of interaction audio/video

We plan to develop a deep learning algorithm that takes a video sequence as input and can predict corresponding sound features with a trained neural network. The sound features are then matched to a database of impact sounds features, and the best matches are stitched together and played back to the user.

We expect this method to be difficult to implement, but yield realistic results and easier to generalize to different interactions with objects.

3.3 Experiment

We propose a psychophysical experiment to evaluate the following hypotheses:

Hypothesis 1 *With varying object density and matching (real) sound, object density will be determined correctly (i.e., correctness will not significantly differ from 100%).*

Hypothesis 2 *With varying object density and fixed neutral sound (white noise), object density has to be guessed (i.e., correctness will not significantly differ from 50%).*

Hypothesis 3 *With fixed object density and varying (pre-recorded) sounds, object density will be determined based on the sound, independent of the actual density.*

The platform to be used will consist of real 3D printed cubes of solid and hollow density firmly attached to the surface of a desk. There is no visual indication that the cubes are any different. A piezoelectric microphone is attached to the side of the cube, and the user wears external sound damping headphones. The dependent variable is *Perceived density* and can be either hollow or solid. Independent variables are *Cube density* (hollow or solid) that is changed throughout every iteration of the experiment, and *Interaction audio* which can be real sound, fixed sound (static noise), augmented hollow sound or augmented solid sound.

Procedure

A cube of variable density is placed in a holder on a table surface. Participants sit in front of the cube and can see it throughout the experiment. Participants are asked to only tap the cube on the top side, with their index finger. They are asked to decide whether the cube is hollow or solid, and until they do they can tap as many times as desired. After the participant decides this, the cube is changed with a cube with different object density (solid or hollow). Participants will not hear or see this exchange. In a first condition, verifying Hypothesis 1, participants hear real sound emitted by the cube. In a second condition, verifying Hypothesis 2, participants wear headphones that also cancel external sounds and hear fixed audio (static noise). In a third condition, verifying Hypothesis 3, participants wear headphones and hear augmented contact sound generated by our system. The augmented sound resembles either hollow or solid audio, and varies throughout every iteration of the procedure.

4. CONCLUSION

In this position paper we discussed an AR system that can manipulate perceived object physical properties. Three methods are proposed to simulate object density by augmenting interaction audio.

We expect that the findings of this research give insight into the effect of audio cues on our multimodal perception of object density of real objects. Furthermore the methods described can be used to simulate density of real objects, to an extent. 3D printed objects can be perceived as having a desired range of densities, without having to change the actual model or material, saving both time and material costs. Furthermore, this research contributes towards the ultimate goal of changing perception of physical properties to simulate all possible real objects.

Finally, for future work we introduce visual cues of object density to the AR system. We then explore the effect of augmented interaction sound on users' ability to discriminate between hollow and solid objects.

5. ADDITIONAL AUTHORS

Additional authors: Nina Rosa (Utrecht University, email: `n.e.rosa@uu.nl`).

6. REFERENCES

[1] A. Charpentier. Experimental Study of Some Aspects of Weight Perception. *Archives de Physiologie Normales Pathologiques*, 3:122–135, 1981.

[2] D. E. DiFranco, G. L. Beauregard, and M. A. Srinivasan. The Effect of Auditory Cues on the Haptic Perception of Stiffness in Virtual Environments. In *Proceedings of the ASME Dynamic Systems and Control Division*, volume 61, pages 17–22, 1997.

[3] T. Hachisu, M. Sato, S. Fukushima, and H. Kajimoto. Augmentation of Material Property by Modulating Vibration Resulting from Tapping. In *Proceedings of the International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 173–180, 2012.

[4] S. Hashiguchi, Y. Sano, F. Shibata, and A. Kimura. RV Dynamics Illusion: Psychophysical Influence on Sense of Weight by Mixed-Reality Visual Stimulation of Moving Objects. In *Proceedings of the International Conference on Virtual, Augmented and Mixed Reality*, pages 55–64, 2014.

[5] Y. Hirano, A. Kimura, F. Shibata, and H. Tamura. Psychophysical Influence of Mixed-Reality Visual Stimulation on Sense of Hardness. In *Proceedings of IEEE Virtual Reality*, pages 51–54, 2011.

[6] M. Kagimoto, A. Kimura, F. Shibata, and H. Tamura. Analysis of Tactual Impression by Audio and Visual Stimulation for User Interface Design in Mixed Reality Environment. In *Proceedings of the International Conference on Virtual and Mixed Reality*, pages 326–335, 2009.

[7] T. Narumi, S. Nishizaka, T. Kajinami, T. Tanikawa, and M. Hirose. Augmented Reality Flavors: Gustatory Display Based on Edible Marker and Cross-modal Interaction. In *Proceedings of Human Factors in Computing Systems*, pages 93–102. ACM, 2011.

[8] P. Punpongsanon, D. Iwai, and K. Sato. SoftAR: Visually Manipulating Haptic Softness Perception in Spatial Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1279–1288, 2015.

[9] N. Rosa, W. Hürst, W. Vos, and P. Werkhoven. The Influence of Visual Cues on Passive Tactile Sensations in a Multimodal Immersive Virtual Environment. In *Proceedings of the ACM International Conference on Multimodal Interaction*, pages 327–334, 2015.

[10] K. Van Den Doel, P. G. Kry, and D. K. Pai. FoleyAutomatic: Physically-based Sound Effects for Interactive Simulation and Animation. In *Proceedings of ACM Computer Graphics and Interactive Techniques*, pages 537–544, 2001.

B. **Experiment 2 Info Sheet**

The total duration of this study will be around 30 minutes and there is no financial compensation.

This study investigates the effect sound and vibro-tactile feedback (vibrations) have on our ability to perceive several objects properties. Amongst others, we can explore and estimate how *thick* an object is by touching it. Your task will be to try and determine the thinner of two objects in front of you by tapping on them with your index finger. You will be using noise canceling earmuffs in combination with in-ear earplugs at different points of the study:

- The in-ear earplugs are used to play audio, and are cleaned before use of every participant.

- The earmuffs prevent external sounds and are cleaned before use of every participant. The earmuffs are slightly tight on the head and therefore we encourage you to let us know if you feel any discomfort like dizziness or fatigue during the experiment.

We will kick off this experiment by going through a training round. This is so that you can get familiar with the task that you will be repeating and perform this task in the same way every time. During this experiment, you are always answering the question: “Which plate in front of you is thinner, the *left* or the *right* one?” When told to do so by the examiner you tap on the finger-shaped indications of the left plate *three times*, then on the right plate three times. After you indicate which plate you think is thinner by saying “left” or “right”, you rotate your chair so that you are facing the wall behind you. During this time the examiner will change the thicknesses of the plates, and will notify you when you can rotate back to repeat the task.

There are some points to note before we start:
- If you have questions at any point, you can ask the examiner.
- Please only *tap* the finger-shaped point on the plates, and only when the examiner gives you the thumbs-up. *No other form of interaction is allowed*.
- If you feel any discomfort at any time, please let us know immediately, as we might have to stop the study.
- If you, at any moment, want to stop the experiment, please say so. You will not be asked to provide a reason.