TU Delft

Delft University of Technology

Visual communication of how fabrics feel

Wijntjes, Maarten W.A.; Xiao, Bei; Volcic, Robert

DOI
10.1167/19.2.4

Publication date
2019

Document Version
Final published version

Published in
Journal of vision

Citation (APA)
Wijntjes, M. W. A., Xiao, B., & Volcic, R. (2019). Visual communication of how fabrics feel. *Journal of vision*, *19*(2), 1-11. https://doi.org/10.1167/19.2.4

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Visual communication of how fabrics feel

Maarten W. A. Wijntjes Delft University of Technology, Delft, the Netherlands

Bei Xiao Department of Computer Science, American University, Washington, DC, USA

Robert Volcic Department of Psychology, New York University Abu Dhabi, Abu Dhabi, UAE

Although product photos and movies are abundantly present in online shopping environments, little is known about how much of the real product experience they capture. While previous studies have shown that movies or interactive imagery give users the impression that these communication forms are more effective, there are no studies addressing this issue quantitatively. We used nine different samples of jeans, because in general fabrics represent a large and interesting product category and specifically because jeans can visually be rather similar while haptically be rather different. In the first experiment we let observers match a haptic stimulus to a visual representation and found that movies were more informative about how objects would feel than photos. In a second experiment we wanted to confirm this finding by using a different experimental paradigm that we deemed a better general paradigm for future studies on this topic: correlations of pairwise similarity ratings. However, the beneficial effect of the movies was absent when using this new paradigm. In the third experiment we investigated this issue by letting people visually observe other people in making haptic similarity judgments. Here, we did find a significant correlation between haptic and visual data. Together, the three experiments suggest that there is a small but significant effect of movies over photos (Experiment 1) but at the same time a significant difference between visual representations and visually perceiving products in reality (Experiments 2 and 3). This finding suggests a substantial theoretical potential for decreasing the gap between virtual and real product presentation.

Introduction

When we see an object, we generally have a good idea of what it would feel like (Baumgartner, Wiebel, & Gegenfurtner, 2013; Bergmann Tiest & Kappers, 2007; Xiao, Bi, Jia, Wei, & Adelson, 2016). This visual prediction of haptic material properties has obvious advantages for real world object interaction, but also practical implications for retail companies that rely on visual communication. If humans would be completely incapable of predicting how products feel on the basis of visual information, then the success of online shopping would only be a fraction of what it is today. Any visual representation of an object that reveals just a little more than only its color will reveal at least something of what it would feel like. But how much can be predicted from a certain form of visual representation? How much useful information does a photo, a video, or an (interactive) animation contain about the haptic material properties? These questions appear key to the online shopping applications, but are also fundamental for the understanding of cross-modal perception.

Bergmann Tiest, and Kappers (2007) showed that correlations between visual and haptic estimates of roughness are relatively high and of the same order as correlations with various physical parameters of roughness. Interobserver correlations within modalities were stronger than correlations between modalities indicating that there are some modality specific differences in roughness estimation that are nonrandom. Additional evidence that vision and touch estimate material properties in a similar fashion comes from Baumgartner et al. (2013). Besides roughness estimates, they considered eight other properties (such as glossiness and hardness) and also included a categorization task. Correlations between modalities were strong for each of the nine material properties, but categorization differed somewhat between the senses: Haptics performed worse than vision.

Connecting visual appearance to mechanical and tactile properties (such as stiffness and heaviness) is especially important in automatic recognition of properties of fabrics. In order to plan actions, a robot

Citation: Wijntjes, M. W. A., Xiao, B., & Volcic, R. (2019). Visual communication of how fabrics feel. *Journal of Vision*, *19*(2):4, 1–11, https://doi.org/10.1167/19.2.4.

https://doi.org/10.1167/19.2.4 Received May 1, 2018; published February 6, 2019 ISSN 1534-7362 Copyright 2019 The Authors

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Downloaded from jov.arvojournals.org on 02/22/2019

usually uses vision to estimate how a piece of cloth would feel before touching. Recent work in computer vision developed algorithms that can automatically estimate mechanical properties of cloth from images, and the consensus is that information extracted from videos is more robust than those from still images. Bouman, Xiao, Battaglia, and Freeman (2013) showed that the human observers' estimation of cloth stiffness and mass were well correlated with the log-adjusted physical parameter values when the video stimuli were presented. Yang, Liang, and Lin (2017) used deep learning neural network and combined the appearance and motion information to classify cloth. Most recently, Bi, Jin, Nienborg, and Xiao (2018) predicted human perception of cloth stiffness using dense motion trajectories extracted from videos. In the multisensory domain, Yuan, Wang, Dong, and Adelson (2017) trained Convolutional Neural Networks to match the visual information such as color and depth images of draping fabrics to the tactile information captured by Gelsight tactile sensor (Johnson & Adelson, 2009). They also showed that a system jointly trained on vision and touch data can outperform a similar system trained only on visual data when tested purely with visual inputs, confirming the importance of multisensory inputs on estimation of fabric properties.

In human vision, a more direct assessment of how vision predicts haptics was used by Xiao et al. (2016). This appears to be the first study to focus on how different visualization styles affect visual and haptic matching of materials. Using various fabrics in a visual-haptic match to sample task, they found that both color and 3D (draped vs. flat) information of the images significantly improved performance. Note that this match-to-sample task directly quantifies the predictive power of a visual representation. The current study uses this experimental paradigm as a starting point to investigate the predictive strength of various photos and movies of fabrics. However, before zooming in on our study, we first discuss the applied context: the domain of online shopping.

Marketing research has acknowledged relatively early (with respect to the age of online shopping) the potentially problematic absence of haptic information in online shopping (Citrin, Stem, Spangenberg, & Clark, 2003). At roughly the same time, the potential of interactive information on online shopping was identified (Childers, Carr, Joann, & Carson, 2001). The importance of haptic information for product evaluation was shown to be strong but also characterized by individual differences (Peck & Childers, 2003). While in subsequent years evidence for the importance of touching products steadily grew (Peck, Barger, & Webb, 2013; Peck & Wiggins, 2006), actual interactive graphics started to become available. Padilla and Chantler (2011) designed an interface called "ShoogleIt" that essentially lets a user scroll through a movie file with a swiping movement on a touch screen. When the movie (or image sequence) is cleverly shot, the swipe movement mimics actual physical interaction with a cloth (Atkinson et al., 2013). Furthermore, Atkinson et al. (2013) showed that estimates of four material attributes (roughness, thickness, elasticity, and temperature) correlated significantly between real touch and virtual interactive touch. Although it is promising that visually mediated haptic information relates to real haptic experience, the study relied crucially on a certain attribute system and more importantly did not include a baseline condition with a simple static photo. Therefore, it is difficult to infer what the added value of interactivity is. A different study on the usefulness of interactive graphics showed that users *think* that interactive graphics (in this case a "Shoogle") give more information about a textile (Overmars & Poels, 2015), called "perceived diagnosticity." But a visualization that makes users *believe* that the information is veridical is something different than a visualization minimizing the difference between visual prediction and actual haptic sensation. Other studies showed that interactive communication influences engagement (Blazquez Cano, Perry, Ashman, & Waite, 2016) or lets users believe the product can really be touched ("It seemed like I could touch. . ."; Verhagen, Vonkeman, Feldberg, & Verhagen, 2014). Whereas convincing graphics may certainly increase sales, they do not solve the fundamental problem of actually being predictive. A product may look appealing, and the consumer may feel like having a correct impression (i.e., prediction), but when the real product does not match this prediction, the visual communication has obviously failed.

It appears that relatively many studies concerning visual product communication in online shopping are concerned with metaperception (reflections/thoughts about sensations) rather than straightforwardly testing the sensorial effectiveness. The reasons could be that it is relatively cumbersome to test the sensorial predictive performance. A match-to-sample task as used by Xiao et al. (2016) is prone to ceiling and/or flooring effects. For example, stimulus samples can be too easily differentiable because of features that have little to do with the material properties, resulting in 100% correct scores for even the lowest quality communication (e.g., low-resolution black and white image). A match-to-sample task is inherently sensitive to features giving away the identity of a stimulus. The stimulus set we used comprised of various jeans fabrics. Jeans generally look rather similar, but can feel very different. Therefore, it is an interesting stimulus material that is not too easy to visually identify, but also not too difficult.

Downloaded from jov.arvojournals.org on 02/22/2019

The main goal of the current study is to understand whether movies better communicate tactile properties than still images. Movement allows the communication of more information and disambiguates the perception of material properties in both human vision (Doerschner et al., 2011; Wendt, Faul, Ekroll, & Mausfeld, 2010) and computer vision (Bouman et al., 2013; Chakravarthi & Pelli, 2011; Yang et al., 2017). Yet, using a movie implies choosing between the many possibilities that could potentially reveal material qualities less visible in a static picture. For example, specific weight could be revealed by a freely moving or falling cloth, the flexibility by spanning and bending the fabric over a bendable shape, etc. To this end, we used six different movie styles to understand what kind of dynamic information communicates the material optimally.

A secondary goal of this study is to understand what experimental paradigms can be used to quantify the effectiveness of a visual communication of tactile properties. As discussed, most related studies have made use of metaperception judgments ("how diagnostic is this presentation?") instead of directly addressing the effectiveness. In the first experiment we used the match-to-sample task as proposed by Xiao et al. (2016) but added a method that corrects wrong matches based on the perceptual distances between the stimuli. The perceptual similarities arising from this method were subsequently used in Experiments 2 and 3 as an alternative to the match-to-sample method.

Experiment 1

In the first experiment we used a match-to-sample-task to measure how well observers can identify a haptic sample that matches a visual stimulus. The main hypothesis is that movies (dynamic) contain more information to perform this task than photos (static). Since we did not a priori know what type of movie would optimally communicate haptic properties, we used six different movie versions.

Observers

Twenty observers (eight females, 12 males; mean age 23.2 years) participated in the first experiment. As will be described later, a between subject design was used resulting in two groups of 10 (each having four female, six males; mean ages 23.3 and 23.1, respectively). They received €10 compensation. Participants provided written consent. The study was approved by the local TU Delft ethics committee and in accordance with the declaration of Helsinki.

Figure 1. Pictures of the nine fabrics used in the experiments. The picture was the final frame of movie Style 4, i.e., top view of a crumbled cloth. The bottom three pictures were not used as visual stimuli (but were available to touch) in Experiment 1.

Stimuli

Nine jeans samples were used as stimuli, shown in Figure 1. They are all considered "jeans," but they varied considerably in smoothness, compliance, elasticity, and weight. We did not measure any of these attributes in terms of physical parameters. Furthermore, they varied in color and 2D textures, which is obviously not of interest to the haptic modality but could possibly influence visual judgments.

During the experiment, the fabrics hung in a box side by side and were obscured to vision by a sheet of white cotton (see Figure 2). Holes allowed for manual exploration.The cloth samples were filmed in the following six different ways:

Style 1: Cloth was stretched over a cylindrical foam shape that was bent and straightened during the movie (roughly resembling a limb joint rotation).
Style 2: The cloth hung down in midair while in the middle of the bottom side a wire was attached. During movie recording, this wire was lifted and then released, so that the cloth fell down.
Style 3: A close-up was made of the texture of a folded cloth that was moved with the hands outside the field of view.
Style 4: Two (female) hands touched and wrinkled the cloth. At the end of the wrinkling, the hands released the cloth that unwrinkled into some static equilibrium position. The movie was taken from an overhead camera viewpoint (pointing downwards to the table).
Style 5: This was roughly similar to Style 4, except that the viewpoint was more "first person," having the camera at chest height.
Style 6: Cloth was draped over a sphere, while the camera was attached to a light source. The camera/light source made a small (~20°) rotational movement around the cloth.

Downloaded from jov.arvojournals.org on 02/22/2019

Figure 2. Experimental setup. In the wooden box, behind the curtain, nine fabrics hung indicated by the nine labels. Observers could feel the fabrics through the holes. Behind the box, a computer screen displayed the visual stimuli.

Sample frames of the six movies styles are shown in Figure 3. For each of the movies, a movie still was chosen that appeared most informative and did not contain motion blur.

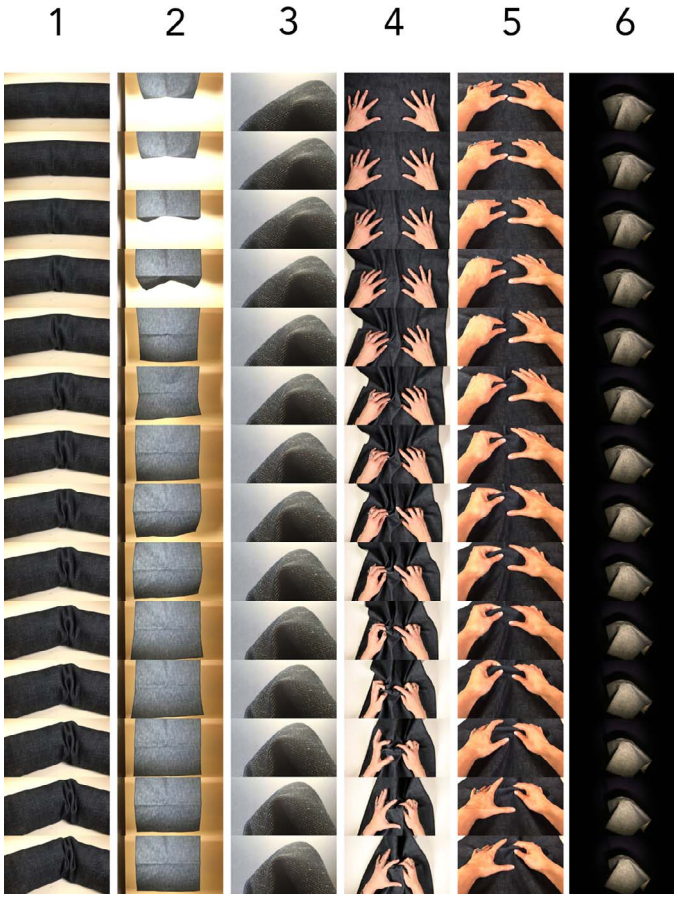

Figure 3. Movie frames of the six different movie styles.

Procedure

The experiment consisted of two parts. In the main part the match-to-sample task was performed, and in the second part haptic similarity judgments were collected to define the perceptual metric between the nine fabrics.

Observers first received a written instruction (together with the consent form) followed by a demonstration of the general procedure. Also, they were allowed to haptically explore the visually obscured (see Figure 2) stimuli before the start of the experiment.

On each trial, observers were shown a visual stimulus (either a movie or photo, depending on the group) and were asked to identify the matching haptic stimulus. Haptic stimuli were randomly ordered and labeled from 1 to 9. Identification took place by selecting one of the nine screen buttons with a track pad, as shown in Figure 2. If observers changed their minds during the experiment, they could go back and change their answer, which occurred once in a while. To avoid getting the last answer for "free" by simply choosing the cloth that was left, we only showed a subset of six different cloths. Thus in each block, observers were visually presented with six cloths (each block the same six cloths) and could choose among nine haptic stimuli. In total, the number of match-to-sample trials amounted to 36 (6 movies/pictures $\times$ 6 styles).

After the six blocks we ran a similarity estimation task. The observer saw two label numbers on the screen that indicated the haptic stimulus pair. The observer was asked to estimate the perceptual dissimilarity on a continuous scale ranging from 1 (same) to 10 (different). Each pair was presented resulting in 36 trials. These data were later linearly rescaled to 1 (same) and 0 (different) to serve as a perceptual metric. Observers varied in the range they used for their similarity judgments, from 72% to 100% of the full scale length, rather evenly distributed with a mean of 88%. Thus, a rather large region of the scale was used.

Data analysis

The haptic similarity data were used as an error metric to transform the raw matching data in perceptual matching data. For example, if sample 2 was shown on the screen, and sample 5 was chosen, and their similarity score was 0.82, then 0.82 was the

Downloaded from jov.arvojournals.org on 02/22/2019

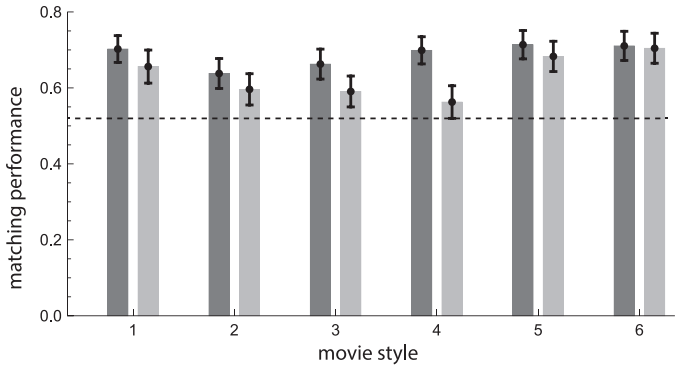

Figure 4. Results of Experiment 1. On the *x* axis, the six different movie styles are shown. Dark bars denote dynamic conditions, light bars static. Error bars denote standard errors of sample mean. The dashed line indicates chance level; see main text for an explanation.

"accuracy" of the match. For matches that were correct, a score of 1 was assigned. The error metric allowed for a fair comparison of judgments because our natural stimuli may not be homogeneously distributed in some conceivable feature space. We used the similarity data on an individual basis; except for four observers we used the mean similarity data (because their similarity data were missing).

Results

The data averaged over 10 observers per condition are shown in Figure 4. To quantify chance level, we took the means of the similarity scores. For a diagonal matrix (in case of uncorrected data, not applying the metric), chance level would be 1/9, but for the actual similarity data (1 = similar, 0 = dissimilar), the particular chance level was 0.52 (indicated by a dashed line in Figure 4). As can be seen, the dynamic presentations (dark gray bars) in all styles result in a higher identification performance than the static presentations. A Mann-Whitney test between the static and dynamic conditions revealed a significant effect ($n1 = n2 = 10$, $U = 77$, $p = 0.023$, one-tailed) with average performance scores of 0.70 (dynamic) and 0.63 (static). To quantify whether this advantage of movies over still images was present on an individual movie style level, we performed separate Mann-Whitney tests for each of these six conditions. Initially, only Style 4 revealed a significant effect, but after correction for multiple comparisons no individual effects were present. To assess performance differences among the various movie styles, we conducted a Friedman test on the dynamic data, averaged over cloths and with movie style as independent variable. This resulted in a nonsignificant effect, $\chi^2(5) = 3.2$, $p = 0.669$, implying that no performance difference was found between the movie styles.We applied the Mann-Whitney tests also to the raw data and found no significant effect of dynamic versus static ($n1 = n2 = 10$, $U = 56$, $p = 0.338$, one-tailed). The absence of a significant effect in the raw data and presence in the (similarity) corrected data motivated us to perform a third test. The main data analysis was performed on data that were corrected by an individual similarity metric: For each (with four exceptions as described in the method section) observer, the similarity data were used as a metric for the matching results. To understand the importance of using the individual metrics, we performed the third Mann-Whitney test on the data that were corrected with the average similarities, i.e., one metric applied to all matching data. Also in this case, we failed to find a significant effect ($n1 = n2 = 360$, $U = 65$, $p = 0.137$) for mean performances of 0.67 (dynamic) and 0.64 (static).

Discussion

Before discussing the main finding, we will briefly discuss the match-to-sample paradigm. The usefulness of a match-to-sample task critically depends on the stimulus set. If the task is too easy (e.g., when clearly distinct fabric categories are used such as leather, canvas, silk, neoprene, etc.), then performance will be at ceiling irrespective of the quality of the visual communication. On the other hand, when the task is too difficult, random responses will also not reveal any interesting difference of our independent variables. As can be seen in our results, the choice of jeans fabrics turned out successful: Neither floor nor ceiling effects mask potential effects within the data. Furthermore, applying the (individual) similarity metric to the raw data increased the sensitivity of our statistics (i.e., they show the effect in the expected direction).

Our hypothesis was confirmed: movies reveal more about how fabrics feel than photos. Observers were better able to identify the haptic fabric samples on the basis of movies than on the basis of still images. The finding does not come as a surprise as it has been previously shown that dynamic information is important for material perception (Doerschner et al., 2011; Wendt et al., 2010). Our finding supports a generalization of these previous findings.

Yet, the beneficial effect was not overwhelming and also only visible when we averaged the data over all six movie styles. On an independent movie level, no performance improvements were found. Furthermore, since we used a still image from the movie, it is possible that the effect is influenced by this choice. Although we did our best in choosing a meaningful still image, we cannot exclude this possibility. Also, a significant effect between images and movies was only found when we used the individual similarity judgments to adjust the

Downloaded from jov.arvojournals.org on 02/22/2019

matching responses. It is difficult to dissociate whether this implies that using individual metrics does indeed reveal an effect more subtle than present in the raw data, or that it randomly affects the data and causes an effect by chance. Yet we believe it is rather plausible that individual differences in haptic fabric perception exist and that they should be taken into account.

To gain more evidence for our hypothesis that movies communicate haptic material properties better than photos, we decided to run a second experiment. We conjectured that the haptic similarity data from Experiment 1 could be compared with visual estimates of haptic similarities. This is a rather different paradigm than the match-to-sample experiment: Instead of analyzing direct comparisons between haptics and vision, we now want to analyze relative differences within modalities and see if these are similar across modalities. If haptic similarity judgments based on movies correlate better with the actual haptic similarities than judgments based on photos, this would strengthen our claim. A second reason for this experiment is that for future studies with a similar question ("which visual representation optimally communicates how products feel?") it could be a more practical experimental paradigm than the match to sample task. For a set of products, baseline haptic similarity judgments can be collected in a local lab setting, whereas a variety of visual explorations can be tested online. There is only one haptic experience, but the possibilities for designing visual representations are infinite. Also, since relative judgments are asked, it could be that floor or ceiling effects play a lesser role than in match-to-sample tasks, although that is not a problem in the current study.

Experiment 2

In the visual equivalent of the haptic similarity judgments from Experiment 1, observers have to judge *whether two samples visually appear to feel similar*. Thus, observers are explicitly instructed to form a prediction of the haptic sensation on the basis of visual information. This yields $n(n-1)/2 = 36$ pairwise comparisons that can be compared between conditions (such as a haptic and visual condition) but also internally: If observers correlate well with each-other, that could imply that perception is unambiguous, and vice versa.

Method

Participants

For the haptic similarity judgments, data were used from the observers of Experiment 1. For the visual dynamic similarity judgments, nine observers participated (two males, seven females; mean age 23 years). For the visual static similarity judgments eight observers participated (four females, four males; mean age 23 years). All gave their written consent and were reimbursed for their participation. Participants provided written consent. The study was approved by the local TU Delft ethics committee and in accordance with the declaration of Helsinki.

Stimuli

All movie/photo styles except number 3 were used in this experiment. The stimuli were presented in pairs on a computer screen. The reason is for excluding Style 3 was pragmatic (we had only six movies and were unable to shoot the remaining three).

Procedure

Observers were instructed to "estimate how similar (a pair of fabrics) would feel" on the basis of what they were visually presented. The experiment was dived in five blocks in which a specific style was presented; block order was counterbalanced (as far as possible). In each block, the observer made 36, $9(9-1)/2$, judgments, thus resulting in a total of 180 trials per observer. A continuous rating scale was used ranging from 1 (same) to 10 (different), with ticks at integer positions.

Results

We compared the data across the three conditions (haptic, movie, and photos). For each pairwise judgment, the average value was computed per condition, across styles. This results in 36 values per condition that are plotted against each other. Results are shown in Figure 5. Correlations between both visual conditions and the haptic condition were low ($r = 0.18$ and $r = 0.13$ for movie and photo conditions, respectively) and not significant ($p = 0.291$ and $p = 0.488$, respectively). The correlation between movie and photo conditions (two different groups of observers) was high and significant ($r = 0.923$, $p < 0.0001$). We also computed correlations between haptic and visual conditions per movie style. All correlations ($N = 10$) were not significant except between haptic and Style 6 movie ($r = 0.330$, $p = 0.0488$). However, when correcting for multiple comparisons, also this correlation failed to be significant.

Discussion

There was no correlation between what was seen, and what was felt. Instead of finding more evidence for

Downloaded from jov.arvojournals.org on 02/22/2019

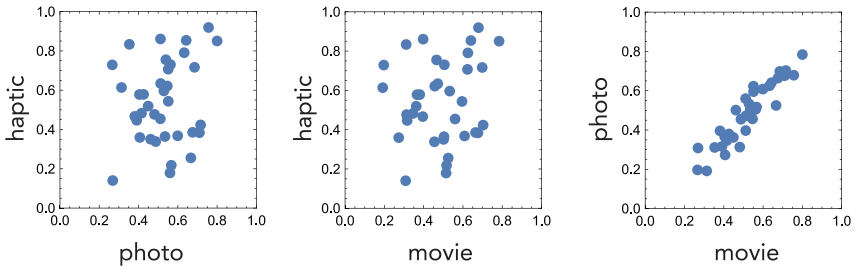

Figure 5. Correlation plots of similarity judgments in Experiment 2 for the three pairs of conditions. As can be seen in the first two plots, there is hardly any correlation between visual and haptic estimates. The last plot shows that visual conditions correlate strongly.

our hypothesis that movies are better than pictures, we found that neither movie nor picture similarity judgments correlates with haptic similarities. On the one hand, this result is surprising because it contrasts with our previous finding. On the other hand, a finding like this is conceivable when considering that the similarity judgments are less direct than match-to-sample estimates. For a match-to-sample estimate, the material feature set of one stimulus is compared to that of another stimulus, across two modalities. Although this process is likely complex (which features are present in both modalities, how are they weighted, etc.), the outcome relies on one comparison step. For similarity judgments, the feature comparison is similar (but maybe less complex because it takes place within modalities), yet the outcome relies on the comparison of the two modality-specific comparisons, i.e., two computations. Thus, the match-to-sample might be less prone to noise (only based on one comparison) than the similarity judgments (comparison between within-vision similarities and those within touch).

If the match-to-sample task is a more direct performance measurement, it would imply that the similarity estimation paradigm is less sensitive in revealing an underlying effect than the match-to-sample paradigm. Another possibility is that observers attend to a different material property set when asked to visually estimate how a stimulus feels than when it is actually touched.

The difference we found in Experiment 1 was relatively small. It could be that the movie quality is partly responsible for the relatively small effect. As previously mentioned, the possibilities for visually representing products are endless. The quality difference between an amateur like the first author and a professional photo/videographer can be substantial. We conjecture that designing a better visual communication would also reveal an effect using the similarity judgment paradigm. Yet, testing all possibilities would be endless.

As an alternative, we chose to take a step back from the production of visual representations, and perform an experiment where observers are seeing (but not touching) the stimuli *in reality*. In Experiment 2 we found no correlation between haptic and visually estimated similarities, which contradicts findings from Experiment 1. If the similarity paradigm is less sensitive, then it could still show an effect when the "signal" is stronger, i.e., when we would have a higher quality visual representation. The obvious candidate for a better visual representation than the pictures and movies we used in Experiments 1 and 2 is clearly reality itself. On the other hand, if the problem is with the similarity paradigm and not with the visual representation, increasing the quality of the visual stimulus would not have an effect.

Therefore, we replicated Experiment 2, but now in a different setting where observers observed other observers interacting with the fabric in a live fashion. There were few restrictions for the noninteracting observers (i.e., they also had audio information) except that they could not actually touch the fabrics.

Experiment 3

Our purpose was to recreate Experiment 2 in reality without mediation of photos and movies. To this end, we designed an experiment with pairs of observers seated opposite to each other. One observer manually explored fabric pairs and was also able to see (and hear) the interaction. The observer seated opposite could see and hear the same (although from an opposite viewpoint) but was not allowed to touch the fabrics.

Participants

Eighteen observers (10 males, eight females) participated in this experiment. The mean age was 32 years.

Stimuli and procedure

We used the same nine fabrics from the previous experiments. As shown in Figure 6, observers were seated opposite each other. The "interacting" participant sat on the left side, the "observing" participant on the right. Before each trial, the experimenter placed two

Downloaded from jov.arvojournals.org on 02/22/2019

Figure 6. Experimental setup of Experiment 3: Two participants sat opposite each other, and both performed similarity judgements on cloth pairs. The "observing" participant was not allowed to touch.

fabrics in front of the interacting participant. He or she was instructed to explore the fabrics to assess the haptic similarity. Observers were explicitly instructed not to communicate with each other, which the experimenter verified through observing them. The "observing" participant was instructed to assess the haptic similarity without touching. Judgments were performed simultaneously. The trackpad of the "interacting" observer was invisible to the "observing" participant. Although this experimental setting allows the participants to see each other, the experimenter did not find any suggestion that they were attending to each other's facial expressions. Verbal communication was forbidden.

Results

The results are shown in Figure 7. The correlation between interacting observers and (purely) haptic similarity data from Experiment 1 was high and significant ($r = 0.871$, $p < 0.0001$). Crucially, the similarity estimates of participants that could only *observe* the other participant interacting with the fabric also correlated significantly ($r = 0.651$, $p < 0.0001$) with the baseline haptic data from Experiment 1. In line with this result, the correlation between observing and interacting was of the same magnitude ($r = 0.622$, $p < 0.0001$).

Discussion

We found a significant relation between nonhaptic (the observation condition) and haptic similarity judgments. This finding contrasts with Experiment 2 in which we relied on the mediating role of our movies; whereas here in Experiment 3 we let participants observe the fabrics directly. Thus, it is possible to infer haptic similarities on the basis of nonhaptic information. The difference between Experiment 1 and 2 was the experimental paradigm, and the difference between Experiment 2 and 3 is the visual representation. The discrepancy between finding a significant difference between movies and pictures in Experiment 1 and not finding this effect in Experiment 2 can thus only be due to a change of paradigm. Yet, Experiment 3 tells us that the similarity paradigm is able to reveal a significant correlation between a haptic and nonhaptic condition. This supports the validity of the similarity paradigm. Together, it suggests that the similarity paradigm is valid, but possibly too weak in revealing an effect when specifically using the visual representations from Experiments 1 and 2. The correlation of similarity judgments between haptic and visual conditions increased from nonsignificant 0.18 and 0.13 for movie and photo conditions, respectively, to a much stronger correlation of 0.65 in Experiment 3. In all experiments the same nine cloths were used, so the correlation increase is very substantial. It shows that the similarity paradigm is able to reveal correlations between modalities.

General discussion

We found that dynamic visual information (a movie) reveals more about how fabrics feel than static visual information (a picture). An attempt to replicate this finding using a different experimental paradigm was unsuccessful: In Experiment 2 we found no facilitating

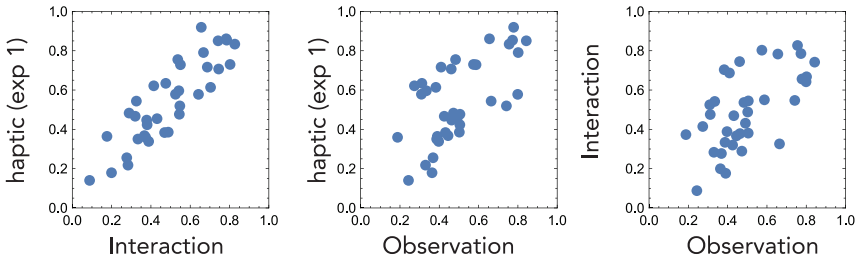

Figure 7. Correlation plots of Experiment 3.

Downloaded from jov.arvojournals.org on 02/22/2019

effect of movies over pictures; we found no correlation between vision and haptics at all. When we directly tested the similarity estimation paradigm on actually seen (but not touched) fabrics in reality, we did find a strong correlation between haptics and vision. Overall, this study shows not only that movies improve visual estimation of haptic material properties (Experiment 1), but also that the type of movie, or communication in general, should be improved to be as realistic as possible (Experiments 2 and 3).

In a related study Xiao et al. (2016) found that visuo-haptic matching improved when color information was present, and when fabrics were draped instead of flattened. Whereas the contribution of color is more of fundamental interest, the influence of draping has practical implications. In the current study, we wanted to continue improving visual representations of fabrics by investigating the role of movement. This "movement" could be a variety of styles, falling, crumbling, bending, or rotating. Although we thought the variety was rather large, the performance results were all very close. Yet, there are many possibilities that may improve the communications. Since in Experiment 3 we found that seeing the fabrics in reality facilitates perceptual performance, but no effect in Experiment 2, there is likely much to improve for our movies. Seeing a product in reality gives "infinite" resolution, no dynamic range problems, no color reproduction problems, perfect binocular disparities, and of course realistic sounds. Although seemingly trivial, it is not very easy to develop a graphics pipeline that copes with each of these facets. On the other hand, if future studies use this reality condition and systematically measure the contribution of the various cues (such as audio or binocular vision), this could lead to a well-informed design that only uses the essential ingredients.

Another reason for the different results of Experiment 2 and 3 could be that the hand movements made in Experiment 3 were actually made by other observers doing the same task. In other words, these were motivated movements whereas the various movements in Experiment 2 were more objective: The movements without hands were all conducted similarly and the hand movements were performed by a person who was not doing any perceptual judgment. In a recent study, Yokosaka, Kuroki, Watanabe, and Nishida (2017) found that visually observing explorative hand movements facilitated material inference. Thus, another way of approaching online product presentations is carefully observing, recording (and maybe transforming) the purposeful interactions people have with real products. This finding could also inspire future computer vision algorithms that estimate material properties from active touching videos. For example, a robot could better learn about fabric haptic properties from watching humans handling the fabric by hands instead of simply watching videos of "what has been done to the fabrics." Indeed, recent work in computer vision shows it is possible to develop algorithms to predict what a person wants to do by observing "hand movements" (Fermüller et al., 2017).

Besides being an interesting solution to the problem of choosing any form of visual representation, using the relatively uncontrolled stimulus presentation of live fabric interaction has certain disadvantages. For example, the interacting participants may spend more time on assessing similar fabric than on dissimilar fabrics. Thus, the observing participant could use exploration time as a cue, which by itself is not an informative cue for haptic fabric properties. We used nine participants interacting with the fabrics and nine other participants observing the fabric interactions. It appears implausible that the interacting participants all followed the pattern of longer time for dissimilar fabrics while the observing participants would at the same time select this pattern as a cue for dissimilarity. In other words, the experimental design likely mitigated the effect of this potential confound. Interestingly, this hypothetical problem would not arise when doing a match to sample task where the visual representation would be replaced by a live interacting participant.

Draping fabrics is already common practice in fabric/product photography, and also movies are sometimes used to display products online. Our results indicate that movies may indeed facilitate perception. There are also other possibilities to represent products online, such as touch screen interactions originally proposed by Padilla and Chantler (2011). Whereas these have been shown to be perceived as diagnostic/informative (Overmars & Poels, 2015), there have been no studies actually testing the effectiveness of these novel visual communication designs. Thus, a natural extension of the study presented here would be to test interactive forms of communication. Our study provides useful paradigms that could facilitate the evaluation of these investigations. Part of the reason we wanted to test the similarity estimation paradigm in Experiment 2 was because of its "scaling" potential: Since the haptic and visual parts are separated, the haptic similarities can be measured in the lab while the visual experiments can be run online.

Understanding the relation between reality and its depiction (whether it is a sketch, painting, photo, computer rendering, hologram, or any future visual invention) is of fundamental interest to broad and diverse audience, from philosophy to the online shopping industry. Every day, people are confronted with a real, physical object that they only knew from its depiction. Although this happens so often, there is little known about the relation between reality/depiction discrepancies and the quality of depicting. The current study contributed to the understanding of this relation

Downloaded from jov.arvojournals.org on 02/22/2019

by zooming in on a particular product category and testing two experimental paradigms and various depictions. It appears promising that movies communicate better how fabrics feel than pictures, but finding the optimal visual representation that captures reality as close as possible remains a future challenge.

Keywords: material perception, fabrics, online shopping, cross-modal transfer, haptics

Acknowledgments

The main funding was provided by the (noncommercial) Netherlands Organization of Scientific Research.

Commercial relationships: Additional funding was provided by the fashion company G-Star, providing the jeans samples, in-kind services and €1.5k cash.
Corresponding author: Maarten W. A. Wijntjes.
Email: m.w.a.wijntjes@tudelft.nl.
Address: Delft University of Technology, Delft, the Netherlands.

References

Atkinson, D., Orzechowski, P., Petreca, B., Bianchi-Berthouze, N., Watkins, P., Baurley, S., ... Chantler, M. (2013). Tactile perceptions of digital textiles: A design research approach. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI '13*, (pp. 1669–1678). ACM. https://doi.org/10.1145/2470654.2466221.

Baumgartner, E., Wiebel, C. B., & Gegenfurtner, K. R. (2013). Visual and haptic representations of material properties. *Multisensory Research*, *26*(5), 429–455, https://doi.org/10.1163/22134808-00002429.

Bergmann Tiest, W. M., & Kappers, A. M. L. (2007). Haptic and visual perception of roughness. *Acta Psychologica*, *124*(2), 177–189, https://doi.org/10.1016/j.actpsy.2006.03.002.

Bi, W., Jin, P., Nienborg, H., & Xiao, B. (2018). Estimating mechanical properties of cloth from videos using dense motion trajectories: Human psychophysics and machine learning. *Journal of Vision*, *18*(5):12, 1–20, https://doi.org/10.1167/18.5.12. [PubMed] [Article]

Blazquez Cano, M., Perry, P., Ashman, R., & Waite, K. (2016). The influence of image interactivity upon user engagement when using mobile touch screens. *Computers in Human Behavior*, *77*, 406–412, https://doi.org/10.1016/j.chb.2017.03.042.

Bouman, K. L., Xiao, B., Battaglia, P., & Freeman, W. T. (2013). Estimating the material properties of fabric from video. *Proceedings of the IEEE International Conference on Computer Vision*, 1984–1991, https://doi.org/10.1109/ICCV.2013.455.

Chakravarthi, R., & Pelli, D. G. (2011). The same binding in contour integration and crowding. *Journal of Vision*, *11*(8):10, 1–12, https://doi.org/10.1167/11.8.10. [PubMed] [Article]

Childers, T. L., Carr, C. L., Joann, P., & Carson, S. (2001). Hedonic and utilitarian motivation for online retail shopping behavior. *Journal of Retailing*, *77*, 511–535, https://doi.org/10.1016/j.jretai.2006.1.

Citrin, A. V., Stem, D. E., Spangenberg, E. R., & Clark, M. J. (2003). Consumer need for tactile input: An internet retailing challenge. *Journal of Business Research*, *56*(11), 915–922, https://doi.org/10.1016/S0148-2963(01)00278-8.

Doerschner, K., Fleming, R. W., Yilmaz, O., Schrater, P. R., Hartung, B., & Kersten, D. (2011). Visual motion and the perception of surface material. *Current Biology*, *21*(23), 2010–2016, https://doi.org/10.1016/j.cub.2011.10.036.

Fermüller, C., Wang, F., Yang, Y., Zampogiannis, K., Zhang, Y., Barranco, F., & Pfeiffer, M. (2017). Prediction of manipulation actions. *International Journal of Computer Vision*, *126*, 358–374, https://doi.org/10.1007/s11263-017-0992-z.

Johnson, M. K., & Adelson, E. H. (2009). Retrographic sensing for the measurement of surface texture and shape. *Computer Vision and Pattern Recognition*, *2009*, 1070–1077.

Overmars, S., & Poels, K. (2015). Online product experiences: The effect of simulating stroking gestures on product understanding and the critical role of user control. *Computers in Human Behavior*, *51*(PA), 272–284, https://doi.org/10.1016/j.chb.2015.04.033.

Padilla, S., & Chantler, M. J. (2011). Shoogleit.com: Engaging online with interactive objects *Digital Engagement*, *2011*(September), 22–25, https://doi.org/10.13140/RG.2.1.1450.1286.

Peck, J., Barger, V. A., & Webb, A. (2013). In search of a surrogate for touch: The effect of haptic imagery on perceived ownership. *Journal of Consumer Psychology*, *23*(2), 189–196, https://doi.org/10.1016/j.jcps.2012.09.001.

Peck, J., & Childers, T. L. (2003). Individual differences in haptic information processing: The "Need for

Downloaded from jov.arvojournals.org on 02/22/2019

Touch" scale. *Journal of Consumer Research*, *30*(3), 430–442, https://doi.org/10.1086/378619.

Peck, J., & Wiggins, J. (2006). It just feels good: Customers' affective response to touch and its influence on persuasion. *Journal of Marketing*, *70*(4), 56–69, https://doi.org/10.1509/jmkg.70.4.56.

Verhagen, T., Vonkeman, C., Feldberg, F., & Verhagen, P. (2014). Present it like it is here: Creating local presence to improve online product experiences. *Computers in Human Behavior*, *39*(September), 270–280, https://doi.org/10.1016/j.chb.2014.07.036.

Wendt, G., Faul, F., Ekroll, V., & Mausfeld, R. (2010). Disparity, motion, and color information improve gloss constancy performance. *Journal of Vision*, *10*(9):7, 1–17, https://doi.org/10.1167/10.9.7. [PubMed] [Article]

Xiao, B., Bi, W., Jia, X., Wei, H., & Adelson, E. H. (2016). Can you see what you feel? Color and folding properties affect visual–tactile material discrimination of fabrics. *Journal of Vision*, *16*(3):34, 1–15, https://doi.org/10.1167/16.3.34. [PubMed] [Article]

Yang, S., Liang, J., & Lin, M. C. (2017). Learning-based cloth material recovery from video. *Proceedings of the IEEE International Conference on Computer Vision*, *201 7*(October), 4393–4403, https://doi.org/10.1109/ICCV.2017.470.

Yokosaka, T., Kuroki, S., Watanabe, J., & Nishida, S. (2018). Estimating tactile perception by observing explorative hand motion of others, *11*(2), 192–203, https://doi.org/10.1109/TOH.2017.2775631.

Yuan, W., Wang, S., Dong, S., & Adelson, E. (2017). Connecting Look and Feel: Associating the visual and tactile properties of physical materials. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5580–5588, https://doi.org/10.1109/CVPR.2017.478.

Downloaded from jov.arvojournals.org on 02/22/2019