Visio-haptic Cross-modal Recognition for Fabrics

ISAMI Chiari\textsuperscript{a}, YAMAMOTO Hiroki\textsuperscript{b}, SUKIGARA Sachiko\textsuperscript{c, *}

\textsuperscript{a} Mechanical Systems Course, Osaka Prefecture University College of Technology, 26-12 Saiwai, Neyagawa, Osaka 572-8572, Japan
\textsuperscript{b} Faculty of Integrated Human Studies, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan
\textsuperscript{c} Department of Advanced Fibro-Science, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan

Received 15 October 2019; accepted for publication 20 November 2019

Abstract

There are many opportunities for visual and tactile interactions with fabrics in daily life, and there have been many studies on visual and tactile evaluation of fabrics. However, there are few studies of how the visual and tactile stimuli influence the recognition of fabrics based on short-term memory. This study examines how the memory of fabric texture is retained and is transferred between visual and haptic perceptions by intra- and inter-modal recognition experiments. Fabric was recognized under both haptic and visual single-modal conditions. Visual recognition was superior to haptic recognition under single-modal conditions, consisting of visual learning followed by a visual test and haptic learning followed by a haptic test. The inter-modal experiment showed that the cross-modal recognition of fabric was possible, regardless of the order of encoding and recalling fabrics by touch or visually. Therefore, single common cross-modal representation is used when fabrics are recognized.

Key Words: Cross-modal, Recognition, Visio-haptic, Fabric texture

1. Introduction

We are constantly in contact with fabric through clothing and feel its texture and hand. Consumers and textile producers use fabric hand evaluation to find suitable end-products; therefore, many studies have examined the texture and hand of fabric. For example, Kawabata and Niwa developed a formula that presents the primary hand values in terms of fabric physical properties [1]. These studies greatly contributed for improving fabric quality. Handling a fabric is a common way to obtain information about it, such as whether it is soft or hard, stretchable or stiff, smooth or rough, and warm or cool. These studies have combined sense-based evaluation with the fabric physical properties to clarify vague information, such as aesthetic perceptions [2, 3], and build a model using artificial intelligence to predict clothing comfort [4]. Thus, most studies have investigated the perception of fabric texture in the moment of seeing or touching a fabric.

However, there are few studies that have focused on memory related to fabric texture. Short-term and working memories are essential for performing complex tasks in daily life [5][6]. For example, short-term memory is used when comparing clothes in e-commerce. Consumers on the internet decide what to purchase based on visual information, such as photographs, and written descriptions, rather than touching the fabrics. This is possible because we can predict tactile texture from visual information to some extent. Thus, the short-term memory of the combination of visual and tactile sensation for fabrics merits investigation.

In this study, we investigate how the memory of fabric texture is maintained for visual and tactile sensation, and across modalities. We clarify which modality is superior for remembering fabric and how fabric texture information is transferred to another modality. This fundamental work could be applied to fabric design improve the first impression of fabrics.

2. Method

2.1 Participants

A total of 42 naïve observers participated in two experiments. Their ages ranged from 20 to 27 years old. All observers reported normal haptic sensation and normal or corrected-to-normal vision. They were divided into two groups depending on learning modality. The first 21 participants (14 men and 7 women) learned by visual sensation, whereas the other 21 participants (12 men and 9 women) learned by haptic sensation.

* Corresponding author: E-mail: sukilaga@kit.ac.jp, Tel: +81-75-724-7365, Fax: +81-75-724-7365
2.2 Fabric stimuli

Stimuli were black woven fabrics comprising nine wool fabrics, a wool/polyester blend, and a wool/polyurethane blend. These stimuli are fabrics which were used for study of aesthetic of textile[7][8].

Table 1 shows the physical properties of the fabrics measured by a Kawabata Evaluation System (KES; Kato Tech, Japan) using the standard conditions. The relationships between the physical properties and human perception are well-known. The stimuli were selected based on bending rigidity (B), sheering rigidity (G), tensile energy (WT), and surface roughness (SMD). The average values for the warp and weft directions are listed in Table 1. The thicknesses and peak values of the heat flux (\(q_{\text{max}}\)), which is the index for the warm or cool feeling of fabrics [9], were similar, except for samples 4–6. Table 2 shows the CIELAB coordinates of the fabric samples. L* (lightness) ranged from 1.66 to 2.70, indicating that the color difference was small.

The size of haptic stimuli was a 20 × 10 cm (warp × weft) piece of fabric and that of the visual stimuli was 6 × 6 cm. For the visual recognition task, the fabric specimen was glued on a ridged surface produced by a 3D printer (Fig. 1) to emphasize the difference in visual material texture.

2.3 Experimental setup

The experimental setup is shown in Fig. 2. A participant sat in front of the stimulus, and their head was immobilized on a chin rest. The visual stimulus was adjusted to a height that was eye-level with the observer 30 cm from the eyes. During the recognition experiment, the visual stimulus was illuminated with a fluorescent light. The illuminance in the horizontal direction in front of the stimulus was 371.2 lx, whereas in the vertical direction it was 1026 lx. The haptic stimulus was presented to the participant’s dominant hand and placed in a box so that the participant could freely handle the sample but not see it. When the stimuli were exchanged by the experimenter, black paper was placed in front of the participant’s eyes to obscure the process from view.
2.4 Procedure

We used a $2 \times 2$ design with repeated measurements, namely, two learning conditions (vision or haptics) and two test conditions (within or across modalities). These conditions are abbreviated as V–V, H–H, V–H, and H–V; for example, V–H represents visual learning followed by a haptic test. Fig. 3 shows the experimental procedure. The six fabrics were randomly selected from the fabrics listed in Table 1. One of six was selected stimuli (called old fabric). Five stimuli (called new fabrics) were distractors. First, the old fabric was presented either visually or haptically, depending on the experimental condition, to be remembered. The stimulus remained until the observer stated that it was remembered. Second, one stimulus was selected from the six fabrics (the old fabric and distractors) and was presented visually or haptically for 8 s for the observer to evaluate whether it was identical to the one remembered. This process was repeated until the six fabrics were each presented twice, resulting in a total of 12 trials. Then, the observer was asked to remember another fabric as the old fabric, and the identification process was repeated. This entire procedure was repeated a total of three times for each participant. Then, the test modality was changed, and the same experiment was carried out again. In total, the participants perform 36 trials taking about 1 h. The test modality order was counterbalanced for each group.

2.5 Analysis

We used generalized linear mixed models (GLMMs) to analyze the relation between the observer’s responses and the experimental conditions. The GLMMs are regression models. In the models, observer’s responses are expressed as various factors called fixed effects and random effects. The responses ($$\text{Resp}$$) in this study can be modeled as
\[ \text{Resp} = aX + \beta Y + \epsilon \]  

(1)

where \( X \) is a matrix of the fixed effects parameters, \( Y \) is a matrix of the random effects parameters, \( a \) is an unknown vector of the fixed effects regression coefficients, \( \beta \) is an unknown vector of the random effects regression coefficients, and \( \epsilon \) is a residual error term assumed to be normally distributed with mean zero and independent of random effects. The presented stimuli (new or old fabric), experimental conditions (V–V, H–H, V–H, or H–V), and interactions between them were used as fixed effects, and the differences among subjects and stimuli were treated as random effects. The model was fitted with probit link and binomial error structure.

We focused on the interaction between the presented stimulus and conditions to compare the recognition performance under the cross-modal conditions with the single-modal conditions. The stimuli are a factor that changes the responses; for example, if the presented stimulus is "old", the response might be "old" and vice versa. If the response is changed by the experimental conditions, the recognition performance depends on the experimental conditions. To find whether the interaction is a significant factor as a fixed effect, we used Akaike's Information Criterion (AIC) to select the optimal model of equation (1). AIC is an indicator proposed by Akaike [10] and is written as

\[ 
\text{AIC} = -2 \ln(L) + 2k 
\]

(2)

where \( L \) is the maximum likelihood and \( k \) is the number of parameters in the model. Using this indicator allows the selection of a model with the minimum necessary variance [11]. The chi squared test was used for each set of conditions to compare the effect of the type of stimuli type. The GLMMs and AIC calculations in this paper used the glmer function in the lme4 package [12] in the R language. The simple effect and interaction for each contrast were tested by using the testInteractions function in the phia package [13].

3. Results

Table 3 shows AIC values for nine models. Model 3a was selected as the best because it had the lowest AIC values. This indicated that the participants’ response was predicted mainly from the effects of stimulus type (Stim), experimental conditions (Cond), and the interaction between them (Stim:Cond). In addition, the random effects were individual participant’s differences (Subj) and differences between stimulus (Item) that affected only the stimulus type. The statistical parameters for the fixed effects (coefficients, statistical significance etc.) of this model are presented in Table 4. We focused on the interaction, which indicated the difference in recognition performance among the conditions.

First, we compared the recognition performance for the single-modal conditions, visual learning followed by a visual test (V–V) versus haptic learning followed by a haptic test (H–H) conditions. The results are shown in Fig. 4. The horizontal axis indicates whether the presented stimulus was old (familiar) or new (unfamiliar). The vertical axis indicates the observer response probabilities. The closer to 1, the greater the probability of a response of "old"; the closer to 0, the greater the probability of a response of "new". Recognition performance for both single-modal conditions was high. The V–V conditions yielded a value

| Models                              | AIC     | Sig. |
|-------------------------------------|---------|------|
| Only main effects                   |         |      |
| 1a Stim                             | 2804.3  |      |
| 1b Cond                             | 3228.0  |      |
| 1c Stim + Cond                      | 2766.7  | ***  |

| Main effects with interaction       |         |      |
| 2a Stim * Cond                      | 2599.0  | ***  |

| Fixed effects + Random effects      |         |      |
| 3a Stim * Cond + (Stim | Subj) + (Stim | Item) | 2531.3  | ***  |
| 3b Stim * Cond + (Cond | Subj) + (Cond | Item) | 2570.7  |      |
| 3c Stim * Cond + (Stim | Subj) + (Cond | Subj) + (Stim | Item) + (Cond | Item) | 2544.6  | ***  |
| 3d Stim * Cond + (Stim + Cond | Subj) + (Stim + Cond | Item) | 2545.9  |      |
| 3e Stim * Cond + (Stim * Cond | Subj) + (Stim * Cond | Item) | 2596.5  |      |

abbreviate
"Stim" means presented two kind of stimuli
"Cond" means experimental condition
Symbol
"A : B" means the interaction between A and B
"A * B" means A + B + A : B
" A | B" means that to treat B as random effects for A
of nearly 1 for the old stimulus presentation and nearly 0 for the new stimulus presentation, indicating nearly perfect performance. The difference of old versus new was statistically confirmed ($\chi^2 = 161.0, p < 0.001$). The H–H conditions also showed a statistically significant positive slope ($\chi^2 = 50.45, p << 0.001$). The comparison of the V–V and H–H conditions indicated inferior performance of haptic recognition compared with visual recognition (V–V vs. H–H: $p < 0.001$).

Next, we compared the recognition performance between the cross-modal conditions, visual learning followed by a haptic test (V–H) versus haptic learning followed by a visual test (H–V). The results are shown in Fig 5. The slopes for the both conditions were positive (V–H: $\chi^2 = 15.20, p < 0.001$; H–V: $\chi^2 = 9.277, p < 0.01$) and virtually identical (V–H vs. H–V: $p = 0.386$), indicating that recognition was possible and comparable for the two cross-modal conditions. The cross-modal recognition performance was lower than the intra-modal performance (V–V vs. V–H: $p < 0.001$; V–V vs. H–V: $p < 0.001$; H–H vs. V–H: $p < 0.01$; H–H vs. H–V: $p < 0.001$).

4. Discussion

The recognition performance for fabric texture using visual information was better than that using haptic information for single-modal conditions. Moreover, recognition was still possible for cross-modal conditions and the cross-modal recognition performance did not depend on which modality was used for remembering or retrieving.

For the single-modal recognitions, haptic recognition was inferior to visual recognition. There are at least two possible reasons for this. The first is the difference between the modalities in the way the stimulus was presented. Newell et al. argued that the learning time needed to be longer for haptic than for visual for shape recognition [14]. In addition, Ernst et al. found that visual and haptic exploration was not balanced across object orientation. Constraining the visual and haptic stimuli improves recognition performance for object shape [15]. Another possibility is that visual recognition ability is inherently higher than haptic visual.

Table 4  Parameters and significance for describing regression model 3a.

| Parameter                      | Coefficient | S.E.  | z     | P     |
|-------------------------------|-------------|-------|-------|-------|
| Intercept                     | -1.25       | 0.116 | -10.8 | < 0.001 |
| Stim old                      | 1.59        | 0.224 | 7.10  | < 0.001 |
| Cond HV                       | 0.463       | 0.0883| 5.24  | < 0.001 |
| Cond VH                       | 0.417       | 0.119 | 3.51  | < 0.001 |
| Cond VV                       | -0.736      | 0.147 | -4.99 | < 0.001 |
| Interaction Stim Old : Cond HV| -0.921      | 0.190 | -4.85 | < 0.001 |
| Interaction Stim Old : Cond VH| -0.734      | 0.219 | -3.35 | < 0.001 |
| Interaction Stim Old : Cond VV| 1.73        | 0.259 | 6.66  | < 0.001 |

Fig. 4 Recognition performance for single-modal recognition, shown as the response to the presented stimuli type (new as distractor and old as learned stimulus) for visual recognition conditions (V–V) and haptic recognition conditions (H–H). Error bars indicate 95% confidence intervals.

Fig. 5 Recognition performance for cross-modal recognition, shown as the response to the presented stimuli type (new as distractor and old as learned stimulus) for visual learning and haptic testing conditions (V–H) and haptic learning and visual testing conditions (H–V). Error bars indicate 95% confidence intervals.
recognition ability. Consistent with this study, Dopjans et al. found that visual recognition for faces was superior to haptic recognition [16].

As far as we know, the present study is the first investigation of the effect of cross-modal recognition for fabric. We found that the information about fabric texture can be transferred between vision and haptics. Interestingly, cross-modal recognition of fabric was possible, regardless of the order of encoding and recalling fabrics by vision or haptics. Therefore, single common cross-modal representation is used when fabrics are recognized. From the result, it implies that, somehow, people can predict touch feeling by only photo (visual information). However, it is still unclear. It is needed more sample variety.

Previous recognition studies on subjects other than fabrics have shown varied results. V–H conditions was superior to H–V conditions for shape orientation [17], whereas the reverse was found for artificial objects [18], and V–H and H–V conditions were the same for faces [6]. The results of face recognition reported by Dopjans et al. were similar to those in the present study [16]. Together with our results, we speculate that the short-term memory for cross-modal recognition can be shared not only for the shape of objects, but also for the surface texture.

5. Conclusion

We investigated memory of fabric texture and how the texture information was transferred between visual and haptic modalities. An old-new recognition task was conducted with a single-modal (V–V, H–H) and cross-modal (V–H, H–V) conditions. For the single-modal conditions, haptic recognition was inferior to visual recognition. For the cross-modal conditions, the information about fabric texture was transferred between vision and haptics. Interestingly, the recognition performance was comparable irrespective of the order in which the information was transferred between vision and haptics. This result implies single common cross-modal representation is used when fabrics are recognized.

6. Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP15H01764.

References

[1] Kawabata S (1980) “The Standardisation and Analysis of Hand Evaluation, 2nd ed”, The Hand Evaluation and Standardization Committee, The Textile Machinery Society of Japan
[2] Brand RH (1964) Textile Research Journal, 34, 791-804. https://doi.org/10.1177/004051756403400909
[3] Phoophat P, Yamamoto H, Sukigara S (2019) The Journal of the Textile Institute, 110, 412-425. https://doi.org/10.1080/00400500.2018.1531460
[4] Wong AS, Li Y, Yeung PKW (2004) Textile Research Journal, 74, 13-19. https://doi.org/10.1177/004051750407400103
[5] Baddeley A (1992) Science, 255, 556-559. https://doi.org/10.1126/science.1736359
[6] Baddeley A (2010) Current Biology, 20, R136-R140. https://doi.org/10.1016/j.cub.2009.12.014
[7] Kitaguchi S, Kumazawa M, Morita H, Endo M, Sato T, Sukigara S (2015) Journal of Textile Engineering, 61, 31-39. https://doi.org/10.4188/jte.61.31
[8] Kitaguchi S, Kumazawa M, Morita H, Endo M, Sato T, Sukigara S (2017) Journal of Textile Engineering, 63, 33-41. https://doi.org/10.4188/jte.63.33
[9] Kawabata S (1984) Sen'i Kikai Gakkaishi (Journal of the Textile Machinery Society of Japan), 37, 130-141. https://doi.org/10.4188/transjitsmj.37.8_T130
[10] Akaiake H (1973) Proceedings of the 2nd International Symposium on Information Theory, 267-281
[11] Symonds M, Moussali A (2011) Behavioral Ecology and Sociobiology, 65, 13-21. https://doi.org/10.1007/s00265-010-1037-6
[12] Bates D, Mächler M, Bolker B, Walker S (2015) Journal of Statistical Software, 67, 1-48. https://doi.org/10.18637/jss.v067.i01
[13] De Rosario-Martinez Helios, et al. (2015) phia: Post-Hoc Interaction Analysis. R package version 0.2-1. https://CRAN.R-project.org/package=phia
[14] Newell FN, Ernst MO, Tjan BS, Bulthoff HH (2001) Psychological Science, 12, 37-42. https://doi.org/10.1111/1467-9280.00307
[15] Ernst MO, Lange C, Newell FN (2007) Canadian Journal of Experimental Psychology, 61, 242-253. https://doi.org/10.1037/cjep2007025
[16] Dopjans L, Wallraven C, Bulthoff C (2009) IEEE Transactions on Haptic, 2, 236-240. https://doi.org/10.1109/TOH.2009.18
[17] Norman JF, Clayton AM, Norman HF, Crabtree CE (2008) Perception, 37, 185-196. https://doi.org/10.1068/p5679
[18] Wismeijer DA, Gegenfurtner KR, Drewing K (2012) Frontiers in Integrative Neuroscience, 6, 105. https://doi.org/10.3389/fnint.2012.00105