To cite this version :

Stanley TARNG, Yaoping HU, Aïda ERFANIAN, Frédéric MERIENNE - An exploration on the integration of vibrotactile and force cues for 3D interactive tasks - In: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces, Allemagne, 2018-03-18 - IEEE Virtual Reality - 2018

An Exploration on the Integration of Vibrotactile and Force Cues for 3D Interactive Tasks

Stanley Tarng[*] Aida Erfanian[*] Yaoping Hu[*] Frédéric Merienne[†]

[*] Dept. of Electrical and Computer Engineering
Schulich School of Engineering
University of Calgary, CANADA

[†] LE2I UMR6306
Arts et Métiers
CNRS University, FRANCE

ABSTRACT

Vibrotactile and force cues of the haptic modality is increasing used to facilitate interactive tasks in three-dimensional (3D) virtual environments (VE). While maximum likelihood estimation (MLE) explains the integration of multi-sensory cues in many studies, an existing work yielded mean and amplitude mismatches when using MLE to interpret the integration of vibrotactile and force cues. To investigate these mismatches, we proposed mean-shifted MLE and conducted a study of comparing MLE and mean-shift MLE. Mean-shifted MLE shared the same additive assumption of the cues as MLE, but took account mean differences of both cues. In a VE, the study replicated the visual scene, the 3D interactive task, and the cues from the existing work. All human participants in the study were biased to rely on the vibrotactile cue for their task, departing from unbiased reliance towards both cues in the existing work. After validating the replications, we applied MLE and mean-shifted MLE to interpret the integration of the vibrotactile and force cues. Similar to the existing work, MLE failed to explain the mean mismatch. Mean-shifted MLE remedied this mismatch, but maintained the amplitude mismatch. Further examinations revealed that the integration of the vibrotactile and force cues might violate the additive assumption of MLE and mean-shifted MLE. This sheds a light for modeling the integration of vibrotactile and force cues to aid 3D interactive tasks within VEs.

Index Terms: H.5.1 [Multimedia Information Systems]: Artificial, augmented and virtual realities; H.5.2 [User Interfaces]: Haptic I/O

1 INTRODUCTION

In a three-dimensional (3D) virtual environment (VE), information and feedback in the visual modality can be overwhelming to human users [1]. Increasingly, VEs have incorporated feedback of other modalities (such as haptics and auditory) to provide the users alternative feedback for presence and interactivity [1], [4]. Of these modalities, haptics is important because of its relevance to direct manipulation of objects. Two common types of feedback in the haptic modality are through vibrotactile and force cues. A generic mechanism of integrating these cues would be necessary to aid 3D interactive tasks.

A potential candidate of the mechanism is maximum likelihood estimation (MLE) [2], [4]. MLE could explain appropriately the cue integration of different modalities, such as visual and haptic cues [4], force and position cues [2], and auditory and visual cues [7]. MLE gives a prediction of a cue integration based on empirical observations of individual cues [4]. Each observation is a Gaussian distribution, represented by its mean (μ), standard deviation (σ), and amplitude (A). The prediction ($\hat{C}$) of an integration among M individual cues is a summation of the cues' weighted observations ($\hat{C}_i,\ i \in \{1,..,M\}$) as follows [4]:

$$\hat{C} = \sum_{i=1}^{M} w_i \hat{C}_i \ , \quad \text{where} \quad w_i = \frac{1/\sigma_i^2}{\sum_{j=1}^{M} 1/\sigma_j^2} \ . \tag{1}$$

When applying MLE to interpret empirical observations of integrating vibrotactile and force cues, an existing work reported mean and amplitude mismatches [3]. To remedy the mismatches, we propose mean-shifted MLE to take account mean differences between the observations and their cue integration. The mean-shifted weights of vibrotactile and force cues (w_V and w_F, respectively) are formulated as below:

$$w_V = \frac{\mu_F - \mu_{FV}}{\mu_F - \mu_V} \quad \text{and} \quad w_F = \frac{\mu_{FV} - \mu_V}{\mu_F - \mu_V} \ , \tag{2}$$

where μ_V, μ_F, and μ_{FV} are the means of the observations of the vibrotactile cue, force cue and their integration, respectively. We conducted a study to explore the suitability of mean-shifted MLE. In the study, each human participant undertook a 3D interactive task within a 3D VE. The study indicated that mean-shifted MLE remedied the mean mismatch but retained amplitude differences.

2 EMPIRICAL STUDY

Within a VE created with the Unity game engine, the visual scene, the interactive task and the cues of the VE were replicated according to their originals in the existing work [3]. The procedures were also replicated, except all participants in the study were instructed to rely on the vibrotactile cue for the task. This led to a biased reliance on the vibrotactile cue. This biased reliance departed from an unbiased reliance on either of the cues in the existing work [3].

Each participant viewed the visual scene with a pair of 3D shutter glasses along with an IR emitter (nVidia Inc., Santa Clara, USA). For 3D interaction, the participant employed the right hand to hold the stylus of a PHANToM Omni device (Geomagic Inc., USA) on a small table. The visual scene consisted of a high-powered transmission line supported by two towers (60.0 m apart) and located in a mountainous region. The transmission line curved towards the ground due to its weight. The viewport of the VE was on a flying drone, which had a robotic arm equipped with a loop-shaped clamp. The clamp covered the line for detecting defects on the line.

Ten participants (age 27.13 ± 5.13 years old) took part in the study. All participants differed from the counterparts in the existing work and were naïve to the purpose of this study. According to a pre-screening, every participant was right-handed, had intact color vision, and possessed normal or corrected-to-normal vision with stereo acuity of at least 40" of arc. The participant wore an E4 wristband (Empatica Inc., Italy) on the left wrist to monitor his/her physiological signals in real time. An ethics approval was attained at our institute for the study.

---

[*] email: { stanley.tarng, aerfania, huy}@ucalgary.ca

[†] email: frederic.merienne@ensam.eu

IEEE Virtual Reality 2018
18-22 March 2018, Reutlingen, Germany

The participant performed an interactive task. During the task, the participant used the stylus of the Omni device to guide the drone to fly and to move the clamp along the transmission line. Any defect on the line was fed back to the user by a vibrotactile cue and/or a force cue. The vibrotactile cue (200 Hz for 1.0 s) was generated through the first motor of the VibroTac (SENSODRIVE Gmbh, Germany) bracelet. The force cue (0.6 N for 1.0 s) was delivered by the Omni device. There were 5 cue profiles with each profile being one testing block as follows:

- V_co: Only a vibrotactile cue on the right hand, co-located with the Omni stylus.
- V_dis: Only a vibrotactile cue on the right forearm, dis-located from the Omni stylus.
- F_only: Only a force cue delivered by the Omni stylus.
- FV_co: A cue consisted of both V_co and F_only profiles.
- FV_dis: A cue consisted of both V_dis and F_only profiles.

The order of these testing blocks was counter-balanced among all participants. A practice block was present for each participant to learn how to fly the drone and detect defects. A total of 15 randomly located defects were on the transmission line, with every block having different set of defect locations. After every block, the participant answered one cybersickness questionnaire [6] and one perceptual questionnaire [3]. Subjective data of perceived usefulness (*Usf*), effectiveness (*Eff*), pleasure (*Pls*) and workload (*Wld*) were gathered using the perceptual questionnaire; Objective data including task completion time (*TCT*) and percentage accuracy in identifying defects (*Acc*) were collected by the VE application. NASA task load index devised the perceived *Wld* [5]. Length of the procedure for each participant was about 1.75 hours.

3 RESULTS AND DISCUSSION

None of the participants suffered from cybersickness according to the physiological data and the responses to the cybersickness questionnaire. The subjective and objective data were normally distributed and thus eligible for ANOVA and Bonferroni post-hoc analyses. Table 1 depicts all subjective and objective data and their ANOVA outcomes. ANOVA analyses of the subjective data revealed that no significant difference existed among the testing blocks for *Usf*, *Eff* and *Pls*, whereas *Wld* gave a significant difference. The difference arose from the V_co vs. F_only blocks and the V_co vs. FV_co blocks, indicated by Bonferroni post-hoc tests. ANOVA analyses of objective data revealed that no significant difference existed among all testing blocks for *TCT*. However, there was a significant difference among all blocks existed for *Acc*. The F_only block was significantly less accurate than other testing blocks indicated by the Bonferroni post hoc tests. As F_only block is the block without the vibrotactile cue, the vibrotactile cue enhanced the accuracy of detecting defects. All these results corresponded to those in the existing work [3], validating our replication of the visual scene, the interactive task and the cues in the VE. This set the baseline for using MLE and mean-shifted MLE.

Using MLE and mean-shifted MLE, the integration of both cues was examined on *Acc*. A Gaussian distribution was estimated for each testing block. Since the participants were instructed to favor the vibrotactile cue over the force cue, their data were reliance-biased towards the vibrotactile cue. The same estimations were carried out for the unbiased data from the existing work [3].

MLE failed to interpret the observed cue integration based on reliance-biased data in both co-located and dis-located settings, as well as reliance-unbiased data in dis-located setting. These observations were close to those of the vibrotactile only (V_co and V_dis) blocks. In contrast, mean-shifted MLE predictions followed the observations closely. When the force cue was delivered dis-located from the vibrotactile cue, it is plausible that the participants ignored the force cue, making mean-shifted MLE prediction plausible. However, the amplitude remained mismatched between the observations of the cue integration and the mean-shifted MLE predictions. This indicates participants did not completely ignore the force cue when delivered at a co-location with the vibrotactile cue. Compared to MLE, mean-shifted MLE elucidated well the observations of the cue integration in all cases.

Table 1: The means and standard deviations of the subjective data and their ANOVA results among all testing blocks.

Data (objective/ subjective)	**Testing Blocks ($\mu \pm \sigma$)**					**ANOVA**	
	V_ co	**V_ dis**	**F_ only**	**FV_ co**	**FV_ dis**	***$F(4,49)$***	***$p<0.05$***
Usf (%)	67 ± 18	60 ± 22	65 ± 20	65 ± 15	60 ± 24	0.65	—
Eff (%)	63 ± 17	57 ± 19	53 ± 23	61 ± 20	56 ± 23	0.92	—
Pls (%)	65 ± 18	61 ± 22	60 ± 22	69 ± 22	64 ± 23	0.83	—
Wld	126±38	138±26	145±28	144±24	139±33	1.79	✓
Acc (%)	79±12	87±13	35±18	81±11	85±12	24.08	✓
TCT (min)	4.1±0.7	4.1±0.6	4.1±0.5	4.1±0.5	4.5±0.9	1.65	—

To investigate the amplitude mismatch, we plotted the relationship between amplitude differences and the mean-shifted weight difference between the vibrotactile and force cues for all reliance-biased and -unbiased data. This yielded a linear relationship by a least square regression with $R^2 = 0.92$. The relationship explains the amplitude mismatch and indicates that a model without the additive assumption might be necessary.

4 CONCLUSION

We proposed mean-shifted MLE to handle the mean and amplitude mismatches yielded by MLE when integrating vibrotactile and force cues. While mean-shifted MLE remedied the mean mismatch, the amplitude mismatch remained. The amplitude mismatch appeared to have a linear relationship with the difference of mean-shifted weights. Future work is to acquire data for modelling this relationship and elucidating the integration of the vibrotactile and force cues to aid 3D interactive tasks.

REFERENCES

[1] H. Q. Dinh, N. Walker, L. F. Hodges, C. Chang Song, and A. Kobayashi, “Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments,” Proc. IEEE Virtual Real., Houston, TX, USA, Mar. 1999, pp. 222–228

[2] K. Drewing and M. O. Ernst, “Integration of force and position cues for shape perception through active touch,” Brain Res., vol. 1078, no. 1, pp. 92–100, 2006.

[3] A. Erfanian, S. Tarng, Y. Hu, J. Plouzeau, and F. Merrien, “Force and Vibrotactile Integration for 3D User Interaction within Virtual Environments,” Proc. IEEE 3DUI, Los Angeles, CA, Mar. 2017, 87-94

[4] M. O. Ernst and M. S. Banks, “Humans integrate visual and haptic information in a statistically optimal fashion.,” Nature, vol. 415, no. 6870, pp. 429–433, 2002.

[5] S. G. Hart, “Nasa-Task Load Index (NASA-TLX); 20 Years Later,” Proc. Hum. Factors Ergon. Soc. Annu. Meet., San Francisco, CA, USA, Oct. 2006, vol. 50, no. 9, pp. 904–908.

[6] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness,” Int. J. Aviat. Psychol., vol. 3, no. 3, pp. 203–220, 1993.

[7] A. Perez-Bellido, M. O. Ernst, S. Soto-Faraco, and J. Lopez-Moliner, “Visual limitations shape audio-visual integration.,” J. Vision, vol. 15, no. 12, pp. 1-15, 2015.