

Visual-to-Haptic Augmentation in XR: A Wearable Glove for Perceptual Grounding in Multimodal Interaction

Mohd Faisal*, Hamdi Elsaddik, Erhan Baturay Onural, Jihong Zhang, Fedwa Laamarti and A.El Saddik

School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario, Canada

Abstract

Extended Reality (XR) systems increasingly deliver high-fidelity visual and auditory experiences, yet tactile perception remains comparatively underutilized as a modality for enriching embodied interaction. This work presents a visual-to-haptic wearable glove and a feature-based visual-to-haptic mapping algorithm that translates spatial and temporal visual features from images and videos into distributed vibrotactile patterns. The proposed method extracts motion, edge, and brightness cues and fuses them into actuator-level intensity maps aligned with a 29-actuator glove arranged in a five-by-seven layout.

The system is implemented through a modular four-layer architecture comprising the XR environment, media content handling, visual-to-haptic processing, and embedded haptic hardware. A within-subject user study (N = 20) compared visual-only interaction with visual-plus-haptic augmentation across texture-based and dynamic video scenarios. Results indicate that tactile augmentation significantly improves perceived realism in dynamic video scenarios and enhances immersion and visual-tactile correspondence across conditions, with stronger and more consistent effects observed for dynamic visual events.

While the current implementation operates in a single-user, offline-synchronized configuration, the findings demonstrate that vision-driven tactile augmentation can function as a perceptual enhancement layer within multimodal XR systems. Such a layer may provide a foundation for future socially enriched XR environments where coherent multisensory grounding supports higher-level interaction and communication.

Keywords

Extended Reality (XR), Visual-to-Haptic Mapping, Haptic Wearables, Embodied Interaction, Multimodal Augmentation, Perceptual Augmentation

1. Introduction

Extended Reality (XR) systems increasingly support immersive interaction across entertainment, education, training, and collaborative scenarios. While contemporary XR platforms deliver high-fidelity visual and auditory experiences, tactile feedback remains comparatively underrepresented, despite its fundamental role in human perception and embodiment [1, 2]. This imbalance creates a perceptual gap between what users see and what they physically sense, limiting multimodal coherence in immersive environments.

Wearable haptic interfaces, particularly glove-based systems, have been investigated as a means of restoring tactile feedback in XR [3]. Our prior work also explored a vibrotactile glove for texture and dynamic event simulation in VR [4]. Advances in lightweight actuators and textile integration have made vibrotactile gloves practical for immersive applications [5, 6], and tactile cues have been shown to enhance immersion and realism during virtual interaction [1, 7]. However, many existing approaches rely on pre-authored haptic libraries or predefined mappings, restricting adaptability to diverse and dynamic visual content.

Research on multimodal interaction demonstrates that combining visual and tactile cues strengthens embodied perception [8]. In parallel, work on social augmentation in XR highlights the importance of

In Proceedings of the ACM CHI 2026 Workshop on Shaping Future Human Connections: Social Augmentation through XR Technologies, April 13, 2026, Barcelona, Spain.

*Corresponding author.

✉ mmohd055@uottawa.ca (M. Faisal); helsaddi@uottawa.ca (H. Elsaddik); eonur052@uottawa.ca (E. B. Onural); jzhan554@uottawa.ca (J. Zhang); flaamart@uottawa.ca (F. Laamarti); elsaddik@uottawa.ca (A.El Saddik)

id 0000-0003-3363-4758 (M. Faisal)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

enriching perceptual cues to support engagement and inclusivity [9, 10]. While these efforts largely emphasize visual and auditory channels, they ultimately depend on coherent perceptual grounding across modalities. Tactile augmentation therefore represents a complementary pathway for reinforcing multisensory coherence in XR.

Visual-to-haptic rendering provides a mechanism for translating visual information directly into tactile stimulation. Prior studies have mapped image intensity and structural visual features to vibrotactile patterns to simulate surface textures and cross-modal tactile effects [11, 12], yet many systems address static and dynamic content separately, or rely on manually authored effects.

In this work, we propose a feature-based visual-to-haptic mapping algorithm implemented within a wearable glove system that translates spatial and temporal visual features into distributed tactile feedback. The hardware platform extends our earlier haptic glove prototype for immersive VR texture and dynamic event simulation [4], and introduces a unified visual-to-haptic mapping algorithm together with a structured perceptual evaluation framework. Images and videos are converted into actuator-level intensity maps delivered across a 29-actuator glove arranged in a 5×7 layout. The system follows a four-layer architecture comprising the XR environment, media handling, visual-to-haptic processing, and embedded hardware layers.

A controlled within-subject study compared visual-only interaction to visual-plus-haptic augmentation across texture-based and dynamic scenarios. Results indicate that tactile augmentation improves perceived immersion, realism, and visual–tactile correspondence, with stronger effects observed for dynamic events. Although the current implementation operates in a single-user, offline-synchronized configuration, the findings suggest that visual-to-haptic augmentation can serve as a perceptual enhancement layer within multimodal XR systems and may support socially enriched experiences grounded in coherent multisensory interaction.

In this work, perceptual augmentation refers to the systematic enhancement of sensory coherence across modalities, where visual structure is reinforced through aligned tactile feedback. Rather than introducing new semantic or social signals, the proposed system strengthens the perceptual grounding of existing visual content, thereby supporting higher-level interaction in XR.

The remainder of this paper is structured as follows. Section 2 reviews related work. Section 3 presents the system and mapping algorithm. Section 4 describes the experimental evaluation. Section 5 discusses implications and limitations.

2. Related Work

Research on immersive XR increasingly emphasizes multisensory interaction, particularly the integration of tactile feedback alongside visual and auditory channels. Prior work spans wearable haptics, visual-to-haptic rendering, multimodal perception, and socially mediated touch.

2.1. Haptic Wearables and Tactile Augmentation

The sense of touch plays a fundamental role in perception and embodiment, yet remains underrepresented in mainstream XR systems [1, 2]. Wearable haptic devices, especially gloves, have been explored to restore tactile feedback [4, 3]. Advances in lightweight actuators and textile integration have enabled increasingly practical wearable haptic gloves for immersive XR interaction [5, 6]. Pneumatic and electrostatic approaches further expand tactile expressivity [3, 6], while textile-based systems embed vibrotactile elements directly into wearable fabrics to enhance comfort and spatial resolution [13].

Empirical studies report that tactile cues improve perceived immersion and realism [3], and psychophysical studies have shown that vibration amplitude and frequency influence perceived surface roughness and texture granularity [14]. Visual–haptic congruence has been shown to enhance engagement and perceptual coherence in multisensory experiences [8], yet many wearable systems still depend on manually authored vibration libraries, limiting adaptability to dynamic visual content.

2.2. Visual-to-Haptic Rendering

Visual-to-haptic rendering translates image and video content into tactile stimulation. Early approaches mapped pixel intensity and structural image features directly to actuator-level vibration patterns [15]. Later work incorporated visual saliency and temporal feature extraction to improve realism in image- and video-driven haptic systems [16]. Dynamic visual events such as motion and impacts have been encoded into time-varying vibrotactile feedback through multimedia effect generation frameworks [16, 17, 18].

More recently, machine learning methods have been proposed to learn visual-to-tactile mappings using cross-modal generation models [19, 20, 21]. Although these approaches enable complex nonlinear mappings, they typically require large paired datasets and introduce computational overhead that can limit real-time wearable deployment.

Despite these advances, existing systems often treat static texture rendering [14] and dynamic event translation [16, 18] separately, or rely on manually authored mappings [17]. There remains a need for a unified wearable approach that integrates spatial and temporal visual features within a single, interpretable real-time pipeline.

2.3. Multimodal Perception and Social XR Context

Human perception integrates visual and tactile inputs into coherent multisensory experiences. Cross-modal studies demonstrate that synchronized tactile cues enhance engagement and perceived realism [8]. In distributed XR systems, however, latency, jitter, and transmission instability may disrupt perceptual coherence and degrade immersion [22, 23]. Reviews of haptic network protocols further emphasize the importance of ultra-low-latency transmission and adaptive synchronization strategies to preserve multisensory alignment in immersive environments [23]. Pseudo-haptic effects further illustrate how visual cues influence tactile perception, reinforcing the importance of consistent cross-modal grounding [8].

Beyond perceptual integration, recent work in social XR highlights the role of enriched sensory cues in supporting engagement, trust, and mediated interaction [9, 10]. Broader analyses of metaverse-scale XR systems identify multimodal coherence, system integration complexity, and perceptual alignment as central technical challenges for immersive environments [24, 25]. Maintaining consistent cross-modal grounding across visual, auditory, and emerging sensory channels is therefore critical for scalable and socially engaging XR experiences.

Systems for mediated touch and expressive haptics demonstrate how vibrotactile signals can convey affect and social intent [17, 18, 26, 27, 28]. Recent work on digital twin robotic systems further shows that physically grounded haptic interaction enhances social presence in metaverse environments. Remote handshake systems implemented with anthropomorphic robotic arms demonstrate that bidirectional haptic feedback strengthens embodied co-presence and interaction realism [29]. These efforts depend fundamentally on reliable multisensory synchronization and adaptable haptic interfaces.

Despite substantial progress across hardware, rendering algorithms, and social XR research, existing solutions remain fragmented. Many systems focus on either static texture rendering [14] or dynamic event translation [16, 18], or rely on manually authored mappings [17], limiting adaptability. There remains a need for a practical wearable system that integrates visual-to-haptic translation for both spatial and temporal content within a unified XR pipeline. The present work addresses this gap by proposing a feature-based mapping algorithm within a modular glove-based architecture.

3. System Design and Implementation

The proposed system delivers visual-to-haptic augmentation through a modular four-layer architecture: (i) Immersive XR Application Layer, (ii) Media Content Layer, (iii) Visual-to-Haptic Rendering Layer, and (iv) Embedded Haptic Hardware Layer (Figure 1). These layers collectively translate spatial and

temporal visual features into distributed vibrotactile feedback delivered via a wearable glove comprising 29 actuators arranged in a 5×7 spatial layout.

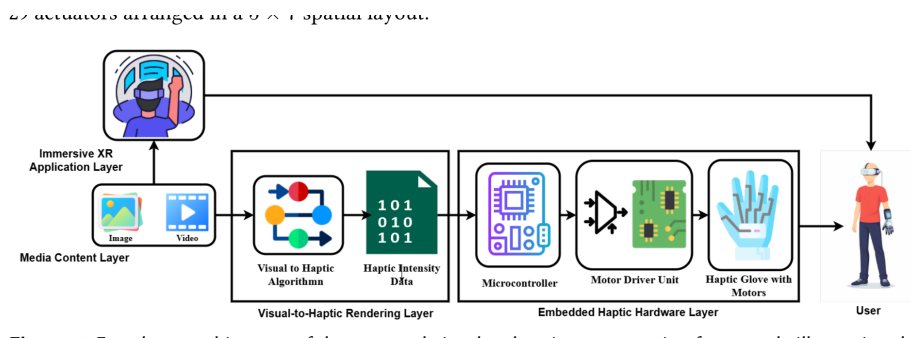


Figure 1: Four-layer architecture of the proposed visual-to-haptic augmentation framework, illustrating the flow from XR visual content to actuator-level tactile output.

The system operates in a frame-based processing mode at 150 ms intervals. This sampling rate was selected to balance perceptual responsiveness with signal stability and actuator limitations. Although currently evaluated in a single-user configuration, the modular structure enables integration into broader XR pipelines, including socially enriched and multi-user environments requiring multisensory grounding.

3.1. Immersive XR Application Layer

This layer generates the visual stimuli that serve as input for tactile translation. The XR environment is implemented in Unity and presents textured virtual objects and embedded video panels designed to evoke tactile interpretations through surface properties and dynamic motion cues (Figure 2).

Visual frames from the rendered scene are streamed to the Media Content Layer for analysis. The current prototype operates in a desktop-based XR configuration; however, the architecture is compatible with VR head-mounted displays and collaborative XR contexts.

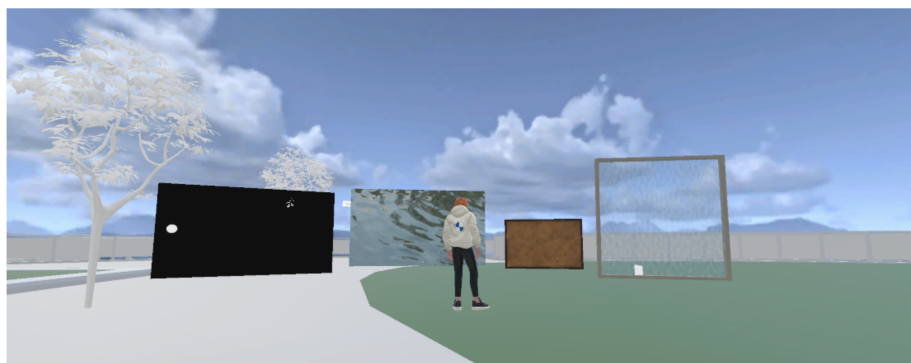


Figure 2: Example XR scene containing textured objects and embedded video panels used to generate visual-to-haptic feedback.

3.2. Media Content Layer

The Media Content Layer manages visual assets embedded within the XR scene, including static images and dynamic video sequences. It forwards raw frames to the Visual-to-Haptic Rendering Layer without modifying visual content.

Two processing pathways are activated depending on input type:

- **Static Image Mode:** A single frame is processed to generate one spatial intensity map.

- **Dynamic Video Mode:** Consecutive frames are sampled at 150 ms intervals to generate temporally indexed tactile sequences.

This separation ensures consistent handling of both spatial textures and time-varying events within a unified rendering framework.

3.3. Visual-to-Haptic Rendering Layer

The Visual-to-Haptic Rendering Layer constitutes the core algorithmic contribution of this work. We propose a feature-based visual-to-haptic mapping algorithm that translates visual structure into actuator-level vibration intensities aligned with the glove geometry.

3.3.1. Grid Mapping and Preprocessing

Each visual frame is segmented into a 5×7 grid corresponding spatially to the 29 actuators on the glove (Figure 3). Six peripheral grid cells are excluded to accommodate ergonomic constraints.

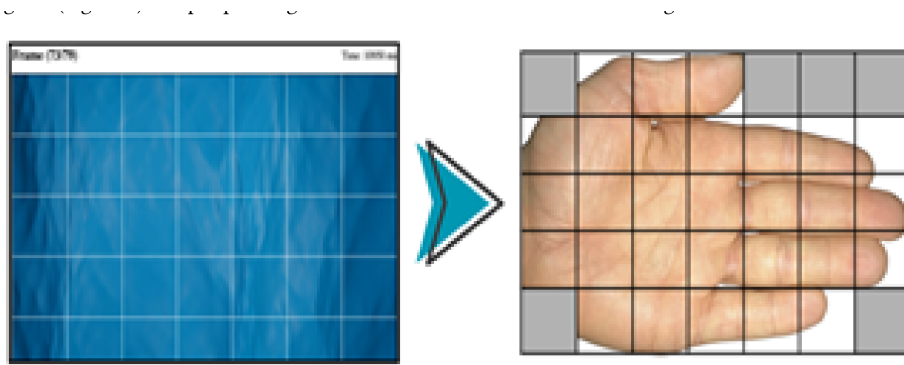


Figure 3: Spatial correspondence between the 5×7 visual grid and the 29-actuator glove layout. Gray cells indicate non-mapped regions.

Before feature extraction, frames undergo:

- Grayscale conversion,
- Gaussian smoothing to reduce high-frequency noise,
- Optional downscaling to ensure computational stability.

These preprocessing steps improve robustness against minor visual fluctuations that could otherwise trigger unintended tactile activation.

3.3.2. Feature Extraction

For each grid cell, three perceptually motivated features are computed:

- **Optical Flow Magnitude (F):** Estimated using the dense Farnebäck method to capture motion intensity between consecutive frames.
- **Sobel Edge Magnitude (S):** Detects structural transitions and surface boundaries.
- **Inverted Brightness (B):** Highlights visually salient darker regions potentially corresponding to texture depth or contrast.

Each feature is normalized to the range $[0, 1]$ to enable consistent fusion across modalities.

3.3.3. Feature Fusion and Parameter Selection

The normalized features are combined using a weighted linear fusion model:

$$V = \alpha F + \beta S + \gamma B$$

where α , β , and γ control the relative contribution of motion, edge structure, and brightness.

Weights were empirically tuned through pilot experiments to balance perceptual saliency and tactile clarity:

- For dynamic video: $(\alpha, \beta, \gamma) = (0.2, 0.4, 0.4)$,
- For static images: $(0, 0.5, 0.5)$.

These weights were selected to balance perceptual saliency with actuator stability and hardware constraints, ensuring clear tactile differentiation without excessive vibration saturation. These values emphasize structural and luminance cues while maintaining responsiveness to motion in dynamic scenarios.

3.3.4. Postprocessing and Temporal Stabilization

To reduce spurious activations, intensity values below a threshold of 15 (on a 0–255 scale) are suppressed. This threshold was determined through preliminary calibration to remove low-amplitude noise while preserving perceptually meaningful tactile cues.

For video sequences, a three-frame moving average is applied to smooth temporal transitions and reduce abrupt actuator fluctuations.

Each processed frame produces a 29-element intensity array corresponding to actuator commands. Static images yield a single tactile map; video sequences generate time-indexed arrays aligned with frame sampling intervals.

3.3.5. Visualization Dashboard

To support inspection and reproducibility, a lightweight browser-based dashboard was developed to visualize actuator intensity distributions during development and testing (Figure 4). The interface displays the visual frame alongside the corresponding 5×7 intensity grid using a color gradient to represent vibration magnitude. Actuator intensities are visualized using a perceptually ordered color gradient ranging from dark blue (low intensity) to red (high intensity), corresponding to normalized vibration magnitudes mapped from 0 to 255. This visualization aids qualitative inspection of spatial distribution and temporal evolution during calibration but does not alter the underlying rendering algorithm.

This tool facilitates qualitative validation of spatial mapping and temporal consistency during system calibration.

3.4. Embedded Haptic Hardware Layer

The haptic glove is a custom-developed prototype designed by the authors. It integrates 29 Linear Resonant Actuators (LRAs) embedded within stretchable textile to ensure comfort and consistent skin contact.

Each actuator is driven by a DRV2605L haptic driver connected through four I2C multiplexers (addresses 0x70–0x73) to a Teensy 4.1 microcontroller. Firmware written in C++ parses actuator intensity arrays, maps values (0–255) to PWM duty cycles, and updates actuators at 150 ms intervals to maintain alignment with visual frame processing.

Power is supplied via a regulated 5 V, 2 A distribution bus. A compact wrist-mounted enclosure houses the controller and Li-Po battery, with flexible silicone wiring supporting user mobility. The hardware architecture allows scalability through additional multiplexers or higher-density actuator configurations in future iterations.

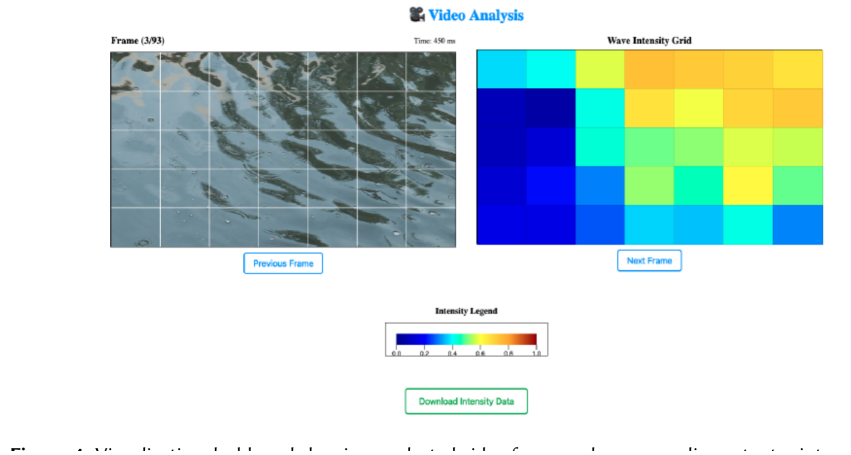


Figure 4: Visualization dashboard showing a selected video frame and corresponding actuator intensity map.

4. Experimental Setup and Evaluation

4.1. Participants and Apparatus

Twenty participants (10 male, 10 female; aged 20–35) voluntarily participated in the study. All reported normal or corrected-to-normal vision and no prior experience with the proposed haptic system. The experimental setup consisted of a *Meta Quest 2* VR headset paired with the custom-developed 29-actuator glove described in Section 3, driven by a *Teensy 4.1* microcontroller and *DRV2605L* haptic drivers.

Experiments were conducted in a quiet indoor environment to minimize external distractions. The study was approved by the University of Ottawa Research Ethics Board (H-09-23-9473). All participants provided written informed consent and responses were anonymized.

4.2. Experimental Design

A within-subject design was adopted to compare visual-only interaction with visual-to-haptic augmentation across two types of XR content. The first module involved interaction with virtual textured surfaces (glass representing smooth texture and sandpaper representing rough texture). The second module involved dynamic video content (flowing water and a bouncing ping-pong ball) designed to emphasize temporal visual features.

Each participant experienced both the visual-only condition and the visual + haptic condition for both modules. The order of conditions was counterbalanced to reduce learning and ordering effects. Each condition lasted approximately two minutes, after which participants completed a structured questionnaire. During haptic trials, tactile signals were synchronized with visual content using the system described in Section 3. Although the sample size ($N = 20$) is typical for within-subject XR studies, future work will include larger cohorts to confirm effect stability.

4.3. Measures and Statistical Analysis

Participants rated their experiences using five-point Likert scales (1 = strongly disagree, 5 = strongly agree). For both modules, perceived realism was collected under visual-only and visual + haptic conditions, enabling direct baseline comparison. Under the haptic condition, additional perceptual dimensions were assessed, including immersion, responsiveness/timing, tactile-visual correspondence, and engagement.

To evaluate the effect of haptic augmentation on realism, paired-sample t-tests were conducted comparing visual-only and visual + haptic conditions. Statistical significance was assessed at $\alpha = 0.05$. Effect sizes were computed using Cohen’s d_z for repeated measures to quantify the magnitude of

Table 1

Mean $\pm$ SD ratings under the visual + haptic condition ($N = 20$).

Metric	Texture	Video
Realism	3.50 ± 1.32	3.95 ± 1.05
Immersion	3.65 ± 1.23	4.10 ± 0.97
Timing / Responsiveness	3.90 ± 1.02	4.30 ± 0.86
Tactile–Visual Correspondence	3.15 ± 1.42	3.75 ± 0.97
Engagement	3.85 ± 1.14	4.30 ± 0.98

observed differences. Descriptive statistics (mean and standard deviation) were computed for all haptic-condition metrics.

4.4. Results

4.4.1. Baseline Comparison: Realism

For the texture module, no significant difference was observed between visual-only ($M = 3.45$, $SD = 1.39$) and visual + haptic ($M = 3.50$, $SD = 1.32$) conditions, $t(19) = 0.11$, $p = 0.914$, $d_z = 0.03$. The negligible effect size indicates that haptic augmentation did not significantly alter perceived realism for static surface textures.

For the video module, haptic feedback significantly increased perceived realism compared to baseline ($M_{visual} = 3.25$, $SD = 1.33$; $M_{haptic} = 3.95$, $SD = 1.05$), $t(19) = 2.15$, $p = 0.044$, $d_z = 0.48$. The moderate effect size suggests that dynamic visual events benefit substantially from tactile augmentation. The 95% confidence interval for the mean difference did not include zero, further supporting the reliability of the observed effect.

4.4.2. Perceptual Ratings Under Haptic Augmentation

Table 1 presents descriptive statistics for perceptual ratings under the visual + haptic condition. Across all perceptual dimensions, the video module yielded higher mean ratings and lower variability, indicating stronger temporal coherence and perceptual consistency when dynamic visual features were translated into tactile signals.

4.5. Qualitative Observations

Open-ended responses reinforced the quantitative findings. Participants frequently described video-based tactile feedback as “natural,” “timely,” and “immersive,” emphasizing strong temporal alignment between visual motion and vibration patterns. In contrast, texture feedback was generally described as distinguishable but occasionally discrete, suggesting that higher actuator density or refined spatial interpolation could further enhance surface continuity.

Overall, the results indicate that the proposed visual-to-haptic rendering approach is particularly effective for dynamic visual content, while static texture rendering may benefit from improved spatial granularity.

5. Discussion

The results indicate that the proposed visual-to-haptic mapping framework effectively translates dynamic visual features into perceptually meaningful tactile feedback in immersive XR. The statistically significant improvement in perceived realism for the video module confirms that motion- and structure-based feature fusion can support coherent cross-modal augmentation when temporal cues are present. The moderate effect size suggests a practically meaningful perceptual benefit for dynamic content.

For static textures, a numerical increase in realism was observed under the visual + haptic condition, though this difference did not reach statistical significance. This suggests that while the current actuator configuration supports perceptible differentiation between smooth and rough surfaces, spatial granularity may limit fine surface continuity. These constraints relate primarily to hardware resolution rather than the feature-based mapping strategy itself. Future iterations will explore higher-density actuator layouts and spatial interpolation strategies to improve continuous texture rendering while maintaining ergonomic feasibility.

Across perceptual metrics, the video module consistently yielded higher ratings and lower variability, indicating stronger temporal coherence when motion-based visual cues were translated into vibration patterns. This supports prior findings in multimodal perception that temporally aligned sensory signals enhance perceptual binding and immersion.

The feature-based mapping approach provides interpretability and low computational overhead compared to data-driven cross-modal synthesis methods, making it suitable for real-time embedded deployment within wearable XR systems.

5.1. Implications and Future Directions

The findings support interpreting visual-to-haptic translation as a perceptual augmentation layer in XR. Rather than introducing new semantic or socially encoded signals, the system reinforces existing visual structure through aligned tactile feedback, strengthening multisensory grounding.

Although evaluated in a single-user configuration, such perceptual grounding may contribute to improved social coordination, communication clarity, and collaborative task performance in multi-user XR environments. Future work will focus on automated event-driven synchronization, optimization of actuator density and spatial distribution, and extended studies incorporating task-based performance measures.

Overall, the proposed framework establishes a practical and interpretable foundation for tactile augmentation in XR, particularly for dynamic visual content.

6. Conclusion

This work presented a feature-based visual-to-haptic glove system capable of rendering both static textures and dynamic visual events in XR. A within-subject study ($N = 20$) showed statistically significant improvements in perceived realism for dynamic video scenarios when haptic feedback was included, while texture differentiation was perceptible but not significantly improved relative to baseline.

These findings validate the proposed mapping framework for dynamic cross-modal augmentation and highlight opportunities for enhanced spatial rendering through increased actuator density. The system establishes a practical foundation for perceptual augmentation in XR, supporting future integration into interactive and socially enriched immersive environments. Such perceptual grounding may contribute to more coherent and inclusive socially augmented XR environments.

7. Acknowledgements

This work has been partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Erhan Baturay Onural is supported by the Republic of Türkiye through the YLSY Study Abroad Programme.

Declaration on Generative AI

The authors have not used any Generative AI tools.

References

- [1] C. Pacchierotti, F. Chinello, K. Koumaditis, M. D. Luca, E. Ofek, O. Georgiou, Guest editorial haptics in the metaverse: Haptic feedback for virtual, augmented, mixed, and extended realities, *IEEE Transactions on Haptics* 17 (2024) 122–128. doi:10.1109/TOH.2024.3369192.
- [2] Y. Tang, J. Xu, Q. Liu, X. Hu, W. Xue, Z. Liu, Z. Lin, H. Lin, Y. Zhang, Z. Zhang, X. Ma, J. Wang, J. Zhong, D. Wang, H. Jiang, Y. Ma, Advancing haptic interfaces for immersive experiences in the metaverse, 2024. doi:10.1016/j.device.2024.100365.
- [3] J. Qi, F. Gao, G. Sun, J. C. Yeo, C. T. Lim, Haptiglove—untethered pneumatic glove for multi-mode haptic feedback in reality–virtuality continuum, *Advanced Science* 10 (2023) 2301044. URL: <https://advanced.onlinelibrary.wiley.com/doi/abs/10.1002/advs.202301044>. doi:<https://doi.org/10.1002/advs.202301044>. arXiv:<https://advanced.onlinelibrary.wiley.com/doi/pdf/10.1002/advs.202301044>.
- [4] M. Faisal, H. Zhu, F. Laamarti, A. El Saddik, A haptic glove for immersive vr texture and dynamic event simulation, in: 2024 IEEE 3rd International Conference on Intelligent Reality (ICIR), 2024, pp. 1–2. doi:10.1109/ICIR64558.2024.10976841.
- [5] J. Park, Y. Lee, S. Cho, A. Choe, J. Yeom, Y. G. Ro, J. Kim, D.-h. Kang, S. Lee, H. Ko, Soft sensors and actuators for wearable human–machine interfaces, *Chemical Reviews* 124 (2024) 1464–1534. URL: <https://doi.org/10.1021/acs.chemrev.3c00356>. doi:10.1021/acs.chemrev.3c00356. arXiv:<https://doi.org/10.1021/acs.chemrev.3c00356>, PMID: 38314694.
- [6] N. Vanichvoranun, H. Lee, S. Kim, S. H. Yoon, Estatig: Wearable haptic feedback with multi-phalanx electrostatic brake for enhanced object perception in vr, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8 (2024). URL: <https://doi.org/10.1145/3678567>. doi:10.1145/3678567.
- [7] N. Magnenat-Thalmann, U. Bonanni, Haptics in virtual reality and multimedia, *IEEE MultiMedia* 13 (2006) 6–11. URL: <https://doi.org/10.1109/MMUL.2006.56>. doi:10.1109/MMUL.2006.56.
- [8] C. Spence, Y. Gao, Enhancing dynamic visual experiences through touch, *Multisensory Research* (2025) 1–36. URL: <https://brill.com/view/journals/msr/aop/article-10.1163-22134808-bja10161/article-10.1163-22134808-bja10161.xml>. doi:10.1163/22134808-bja10161.
- [9] J. D. Hart, T. Piumsomboon, L. Lawrence, G. A. Lee, R. T. Smith, M. Billinghurst, Emotion sharing and augmentation in cooperative virtual reality games, in: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts, CHI PLAY '18 Extended Abstracts*, Association for Computing Machinery, New York, NY, USA, 2018, p. 453–460. URL: <https://doi.org/10.1145/3270316.3271543>. doi:10.1145/3270316.3271543.
- [10] J. D. Hart, T. Piumsomboon, G. Lee, M. Billinghurst, Sharing and augmenting emotion in collaborative mixed reality, in: *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 2018, pp. 212–213. doi:10.1109/ISMAR-Adjunct.2018.00069.
- [11] V. M. Riojas, Image map based haptic rendering of surface textures on 3-d virtual objects-edición única, 2010. URL: <https://api.semanticscholar.org/CorpusID:64428980>.
- [12] G. Cao, J. Jiang, N. Mao, D. Bollegala, M. Li, S. Luo, Vis2hap: Vision-based haptic rendering by cross-modal generation, in: *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 12443–12449. doi:10.1109/ICRA48891.2023.10160373.
- [13] K. Yao, Q. Zhuang, Q. Zhang, J. Zhou, C. K. Yiu, J. Zhang, D. Ye, Y. Yang, K. W. Wong, L. Chow, T. Huang, Y. Qiu, S. Jia, Z. Li, G. Zhao, H. Zhang, J. Zhu, X. Huang, J. Li, Y. Gao, H. Wang, J. Li, Y. Huang, D. Li, B. Zhang, J. Wang, Z. Chen, G. Guo, Z. Zheng, X. Yu, A fully integrated breathable haptic textile, *Science Advances* 10 (2024) eadq9575. URL: <https://www.science.org/doi/abs/10.1126/sciadv.adq9575>. doi:10.1126/sciadv.adq9575. arXiv:<https://www.science.org/doi/pdf/10.1126/sciadv.adq9575>.
- [14] W. Nai, J. Liu, C. Sun, Q. Wang, G. Liu, X. Sun, Vibrotactile feedback rendering of patterned textures using a waveform segment table method, *IEEE Transactions on Haptics* 14 (2021) 849–861. doi:10.1109/TOH.2021.3084304.
- [15] S. Rasool, A. Sourin, Image-driven haptic rendering in virtual environments, in: *2013 International Conference on Cyberworlds*, 2013, pp. 286–293. doi:10.1109/CW.2013.28.

- [16] M. Kim, S. Lee, S. Choi, Saliency-driven real-time video-to-tactile translation, *IEEE Transactions on Haptics* 7 (2014) 394–404. doi:10.1109/TOH.2013.58.
- [17] M. Abdur Rahman, A. Alkhaldi, J. Cha, A. El Saddik, Adding haptic feature to youtube, in: *Proceedings of the 18th ACM International Conference on Multimedia, MM '10*, Association for Computing Machinery, New York, NY, USA, 2010, p. 1643–1646. URL: <https://doi.org/10.1145/1873951.1874310>. doi:10.1145/1873951.1874310.
- [18] S. Nam, M. Chung, H. Kim, E. Kim, T. Kim, Y. Yoo, Automatic generation of multimodal 4d effects for immersive video watching experiences, in: *SIGGRAPH Asia 2024 Technical Communications, SA '24*, Association for Computing Machinery, New York, NY, USA, 2024. URL: <https://doi.org/10.1145/3681758.3698021>. doi:10.1145/3681758.3698021.
- [19] D. Chen, Y. Ding, P. Gao, H. Ni, H. Zeng, J. Liu, A. Song, Multi-dimensional texture haptic cross-modal generation and display method based on bi-mamba network, *IEEE Transactions on Instrumentation and Measurement* (2025) 1–1. doi:10.1109/TIM.2025.3619260.
- [20] S. Cai, K. Zhu, Y. Ban, T. Narumi, Visual-tactile cross-modal data generation using residue-fusion gan with feature-matching and perceptual losses, *IEEE Robotics and Automation Letters* 6 (2021) 7525–7532. doi:10.1109/LRA.2021.3095925.
- [21] Y. Fang, X. Zhang, W. Xu, G. Liu, J. Zhao, Bidirectional visual-tactile cross-modal generation using latent feature space flow model, *Neural Networks* 172 (2024) 106088. URL: <https://www.sciencedirect.com/science/article/pii/S0893608023007499>. doi:10.1016/j.neunet.2023.12.042.
- [22] U. Ragavee, M. Swetha, P. Gopika, L. M. J. M. Saravanapriya, L. D. KM, Hybrid tactile internet framework revolutionizing real-time human-machine interactions with multi-modal technologies, in: *2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)*, 2025, pp. 1811–1818. doi:10.1109/ICMCSI64620.2025.10883481.
- [23] M. Faisal, R. A. Martinez Velazquez, F. Laamarti, H. Al Osman, A. El Saddik, Haptic network protocols: A comprehensive review and directions for next-gen metaverse applications, *ACM Trans. Multimedia Comput. Commun. Appl.* 21 (2025). URL: <https://doi.org/10.1145/3759459>. doi:10.1145/3759459.
- [24] F. Laamarti, M. Faisal, A. E. Saddik, Exploring the potential and technical challenges of metaverse for university education, in: *2024 IEEE 3rd International Conference on Intelligent Reality (ICIR)*, 2024, pp. 1–2. doi:10.1109/ICIR64558.2024.10976880.
- [25] R. Ferdousi, M. Faisal, F. Laamarti, C. Yang, A. E. Saddik, Exploring user perceptions of virtual reality scene design in metaverse learning environments, in: *2024 IEEE International Conference on Consumer Electronics (ICCE)*, 2024, pp. 1–6. doi:10.1109/ICCE59016.2024.10444490.
- [26] V. Shen, T. Rae-Grant, J. Mullenbach, C. Harrison, C. Shultz, Fluid reality: High-resolution, untethered haptic gloves using electroosmotic pump arrays, in: *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, UIST '23*, Association for Computing Machinery, New York, NY, USA, 2023. URL: <https://doi.org/10.1145/3586183.3606771>. doi:10.1145/3586183.3606771.
- [27] T. Muender, M. Bonfert, A. V. Reinschluessel, R. Malaka, T. Döring, Haptic fidelity framework: Defining the factors of realistic haptic feedback for virtual reality, in: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22*, Association for Computing Machinery, New York, NY, USA, 2022. URL: <https://doi.org/10.1145/3491102.3501953>. doi:10.1145/3491102.3501953.
- [28] Y. Abbass, S. Dosen, L. Seminara, M. Valle, Full-hand electrotactile feedback using electronic skin and matrix electrodes for high-bandwidth human-machine interfacing, *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences* 380 (2022). doi:10.1098/rsta.2021.0017.
- [29] M. Faisal, F. Laamarti, A. El Saddik, Digital twin haptic robotic arms: Towards handshakes in the metaverse, *Electronics* 12 (2023). URL: <https://www.mdpi.com/2079-9292/12/12/2603>. doi:10.3390/electronics12122603.