2025 IEEE World Haptics Conference (WHC)
Suwon Convention Center, Suwon, Korea
July 8 ~ 11, 2025

Generative AI-Based Multimodal Authoring Tool for Virtual Objects

Hyunwook Jung, Seoyong Nam, and Yongjae Yoo

***Abstract*—In this demo, we present a multimodal haptic authoring tool that allows users to create and modify haptic properties for virtual 3D objects. Utilizing AI techniques including computer vision and large language models (LLMs), the tool automatically generates and attributes haptic properties such as temperature, stiffness, and texture to the virtual object, as well as their corresponding physics properties such as mass, friction, conductivity, and elasticity. The generated properties are directly presented to the haptic hardware supported by our tool, allowing users to test the effects iteratively. With *end-to-end* support for haptic authoring, our tool would reduce the efforts on designing haptic experiences in virtual environments.**

***Index Terms*—AI-based haptic authoring, virtual reality, multimodality**

I. Introduction

How to create good haptic effects is an important research question that has been studied for over a few decades. Most efforts were made in vibrotactile authoring, including VibViz [1] and HapticGen [2], which utilize multidimensional visualization and LLM-based recommendation techniques, respectively. However, to create non-vibrotactile modalities, especially for 3D objects in virtual environments, there is still a lack of a prominent and effective way. The authoring process relies on labor-intensive trial-and-error testing for setting parameter values. To address this, we suggest an automated, streamlined tool that leverages AI to generate multimodal haptic properties easily, while providing an intuitive interface for real-time testing with hardware.

II. Implementation

We designed a pipeline for multimodal authoring that consists of three stages: (1) property generation, (2) editing and fine-tuning by users, and (3) rendering through hardware (Fig. 1).

AI-driven Property Generation: In the first stage, the tool runs a multi-view rendering for the target 3D object using Cap3D [3], and then runs vision models such as BLIP-2 [4] and Qwen2.5-VL [5] to extract the descriptions of the object. Large Language Models (LLM; GPT-4o) with customized prompting extract properties of different modalities from the image descriptions. Among multiple properties, we set the models to infer temperature (0-100 °C), stiffness level (0-100) for hardness, and vibration characteristics (roughness, hardness, heaviness, crispiness) for surface texture. To render haptic effects in a more convincing way, we also estimated physical properties of the target object, such as mass, drag, static/dynamic friction, and bounciness, using the same model.

This work was supported by a research grant No. RS-2023-00226263.

All authors are with Department of Applied Artificial Intelligence, Hanyang University, South Korea (emails: {hyunwook7120, namsyong0508, yongjaeyoo}@hanyang.ac.kr).

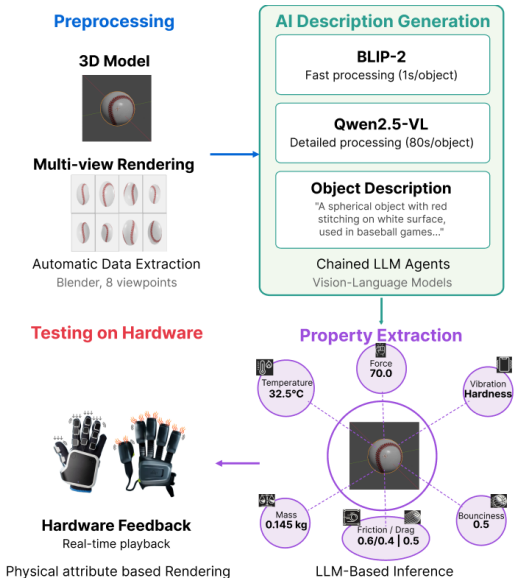

Fig. 1. Multimodal Authoring Scheme of Our AI-based Tool

These values are directly applicable to Unity3D engine's Rigidbody and Material settings.

Parameter Tuning with **Real-time Testing**: The tool allows users to refine these properties with real-time feedback that is presented on the supported hardware. Currently, the tool supports two different exoskeleton gloves (WeArt TouchDiver Pro and SenseGlove Nova 2), but it can be easily expanded by adding another hardware support module. The module translates the properties into device-specific commands, allowing users to experience the generated sensations.

III. Conclusion and Ongoing Works

In this work, we presented an AI-based authoring and testing tool for multimodal haptic effects for virtual 3D objects. Our future work will include improving the AI-based property generation mechanisms and validating the usefulness of the tool via longitudinal user studies.

References

[1] H. Seifi, K. Zhang, and K. E. MacLean. "VibViz: Organizing, visualizing and navigating vibration libraries." in Proc. IEEE World Haptics Conference (WHC), 2015.
[2] Y. Sung, K. John, S.H Yoon, and H. Seifi. "HapticGen: Generative Text-to-Vibration Model for Streamlining Haptic Design", in Proc. CHI Conference on Human Factors in Computing Systems (CHI '25), 2025.
[3] T. Luo, C. Rockwell, H. Lee, and J. Johnson, "Scalable 3D Captioning with Pretrained Models," in Advances in Neural Information Processing Systems 36 (NeurIPS 2023) Datasets and Benchmarks Track, 2023.
[4] J. Li, D. Li, S. Savarese, and S. Hoi, "BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models," in Proc. 40th International Conference on Machine Learning (ICML), 2023.
[5] S. Bai et al., "Qwen2.5-VL Technical Report," 2025.