BEFORE SMELLING THE VIDEO: A TWO-STAGE PIPELINE FOR INTERPRETABLE VIDEO-TO-SCENT PLANS

KAI WANG, KEVIN SHAO, RUIQI CHEN

AUTONOMY, ROBOTICS, & CONNECTIVITY

Background

Smell has strong potential to enhance immersion in interactive media.

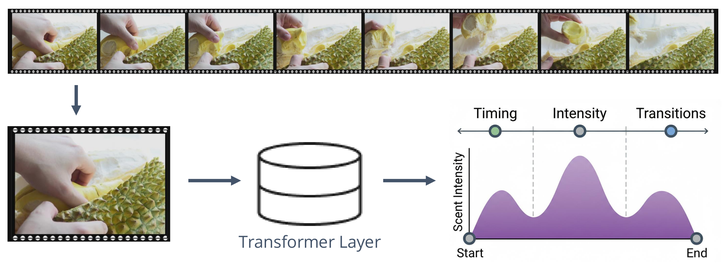
- Olfaction is closely tied to emotion and memory [1].
- Even simple scent cues can increase presence and immersion [2].

However, smell remains rare in deployed video experiences because

- **scent hardware is difficult to control**,
- **scents linger and mix**, and
- **designers lack reliable workflows** for specifying what to smell and when to release it, especially for dynamic video [3] [4] [5].

Research Gaps – Motivations

Existing Work	Research Gaps
Fixed, designer-triggered cues [1] [5].	Limits variation; non-scalable for diverse content.
Optimized for text (Captioning, QA, retrieval) [6].	Lacks physical parameters (intensity, duration).
Static event-to-odor links [1] [7].	Needs a temporal transformation layer for transitions.

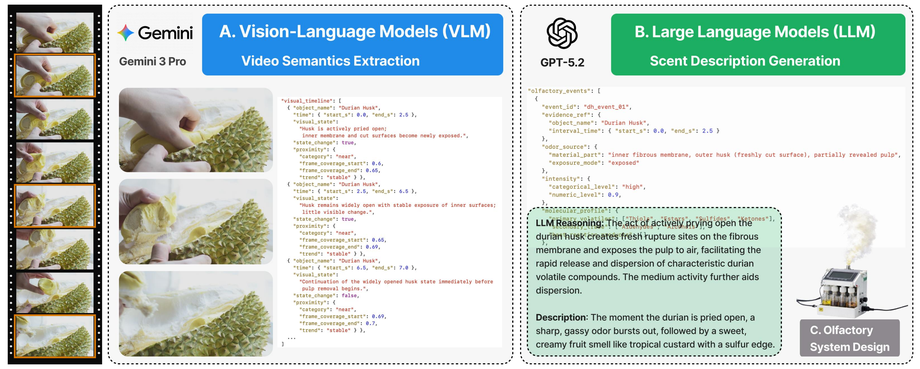


Our work addresses these gaps by enabling more adaptive event-to-odor mapping, parameterized olfactory outputs, and a temporal transformation layer that supports realistic transitions.

References

[1] Brewster et al. 2006. Olfoto: designing a smell-based interaction. In Proceedings of the SIGCHI conference on Human Factors in computing systems. 653–662.
[2] Munyan III et al. 2016. Olfactory stimuli increase presence in virtual environments. PloS one 11, 6 (2016), e0157568.
[3] Jas Brooks and Pedro Lopes. 2023. Smell & paste: Low-fidelity prototyping for olfactory experiences. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 1–16.
[4] Gheorghita Ghinea and Oluwakemi A Ademoye. 2011. Olfaction-enhanced multimedia: perspectives and challenges. Multimedia Tools and Applications 55, 3 (2011), 601–626.
[5] Joseph" Jofish" Kaye. 2004. Making scents: Aromatic output for HCI. interactions 11, 1 (2004), 48–61.
[6] Krishna et al. 2017. Dense-captioning events in videos. In Proceedings of the IEEE international conference on computer vision. 706–715.
[7] Maggioni et al. 2020. SMELL SPACE: mapping out the olfactory design space for novel interactions. ACM Transactions on Computer-Human Interaction (TOCHI) 27, 5 (2020), 1–26.

Two-Stage Video-to-Scent Planning Pipeline

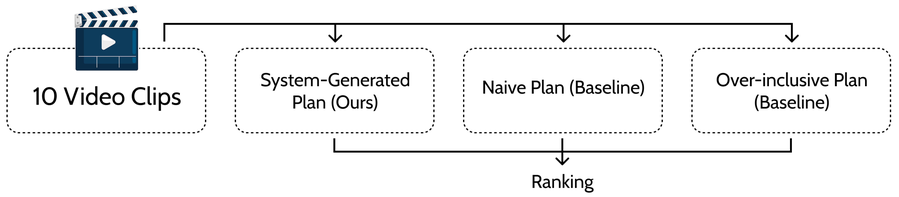


The pipeline is composed of three stages:

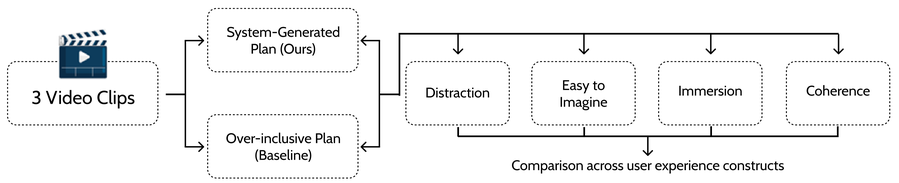
- **Stage A:** VLM (Gemini 3 Pro) extracts time-aligned visual semantics from sampled frames.
- **Stage B:** LLM (GPT-5.2) converts the timeline into a structured scent plan by selecting smell-relevant sources and modeling intensity over time.
- **Stage C:** Output a scent plan for future olfactory interfaces, separating semantic planning from physical scent delivery.

Evaluation Methodology

STUDY 1: Semantic Selection for Olfactory Planning



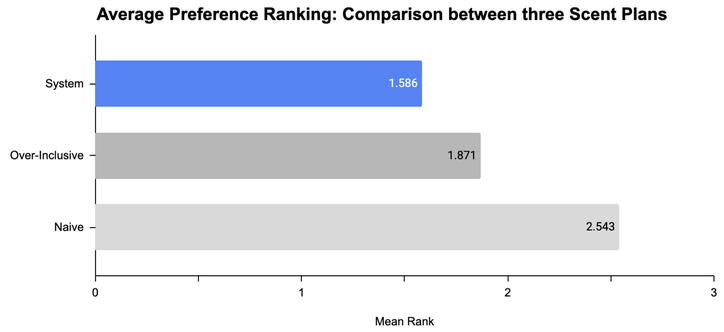
STUDY 2: Experiential Plausibility of System-Generated Plans



Note that three video clips in Study 2 are representative clips selected from Study 1.

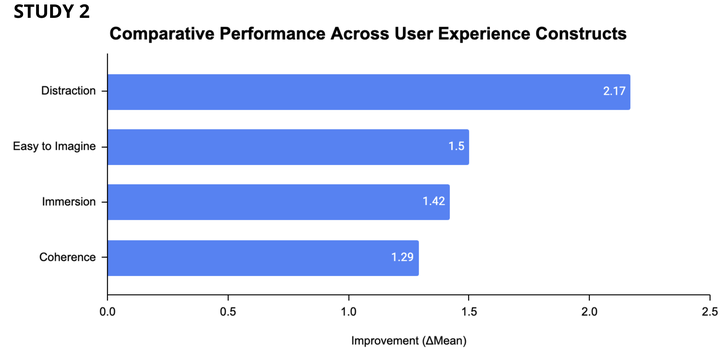
Results

STUDY 1



System-generated plans were the clear favorite. It achieved the best average preference ranking (1.586, where a lower score indicates a higher rank) and was selected as the #1 choice in 54.3% of responses, significantly outperforming both baselines.

STUDY 2



In Study 2, the system-generated plans were perceived as more immersive, more coherent, and easier to imagine than the over-inclusive baseline. All participants (n = 8) found the system plans to be less distracting and more immersive. Positive Δ values indicate a direct advantage for the system over the baseline.

Future Work

This work validates the semantic and experiential intelligibility of scent planning (not end-to-end scent delivery). Future work will

- integrate plans with physical scent-delivery devices,
- add adaptive user controls for timing and intensity, and
- extend to interactive and user-generated media.

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

ADVISORS: DENISE WILSON, SEP MAKHSOUS

SPONSORS: DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING, UNIVERSITY OF WASHINGTON