2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) | 979-8-3503-2838-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISMAR59233.2023.00120

Merging Camera and Object Haptic Motion Effects for Improved 4D Experiences

Jaejun Park *
Interaction Laboratory, POSTECH

Sangyoon Han †
Interaction Laboratory, POSTECH

Seungmoon Choi ‡
Interaction Laboratory, POSTECH

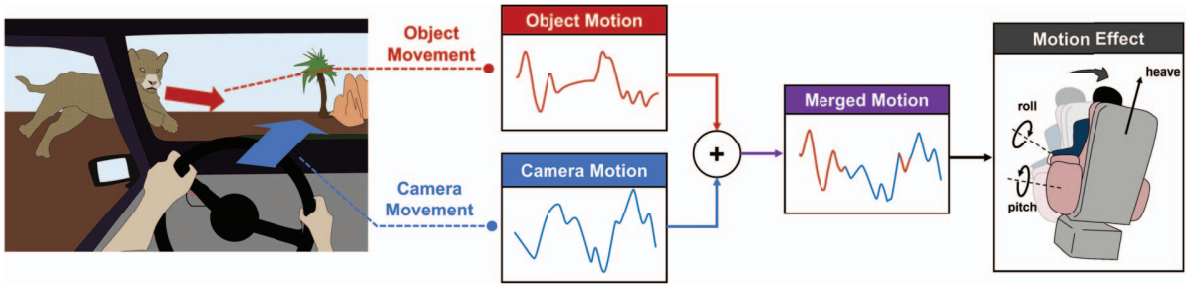


Figure 1: A conceptual overview of merging two haptic motion effects. The virtual camera located between the driver's eyes moves as the car moves, and this camera motion (blue arrow) is made to a camera motion effect for a motion platform. An object motion effect is also generated from the cheetah's motion (red arrow) to emphasize its movement quickly approaching the car. These camera and object motion effects are merged into one final motion effect that describes both effects simultaneously.

Abstract

(Haptic) motion effects refer to the vestibular stimuli generated by a motion platform and delivered to the whole body of a user sitting on the platform. Motion effects are an essential tool for creating vivid sensory experiences in various extended reality (XR) applications, ranging from training simulators to recent 4D rides, films, and games for entertainment. For the latter purpose, motion effects emphasize audiovisual events occurring in the scene, such as camera motion, the movement of an object of interest, and special sounds. Recent research developed several algorithms to produce motion effects from the audiovisual stream automatically. However, these algorithms are designed for a single class of motion effects, and extension to multiple motion effect classes remains unexplored. In this paper, we propose an algorithmic framework that merges camera and object motion effects into one motion effect while preserving the perceptual consequences of the two effects. We validate the framework's perceptual performance through a user study. To our knowledge, this work is one of the first successful reports of merging different kinds of motion effects for improved XR experiences.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction (HCI)—Interaction devices—Haptic devices

1 Introduction

Many applications in extended reality (XR) use vestibular stimuli that convey moving sensations of the user's entire body. Such haptic *motion effects* are part of the standard multisensory experiences, often called 4D experiences, which refer to audiovisual experiences enriched with stimuli in additional sensory modalities such as motion, vibration, wind, heat, and scent. Integrating these sensory elements can bridge the gap between the virtual and physical realms, providing more comprehensive and engaging user experiences [43].

Many motion platforms and applications are available today, including commercial 4D systems for theaters [2, 7, 8] and homes [31–33]. Despite such devices' prevalence, most 4D effects still need to be produced manually, and it is a major bottleneck impeding a widespread distribution of 4D content [27]. In response to these needs, there have been increasing research outcomes regarding the automatic generation of 4D effects from various sources, such as audiovisual content or measured data from physical systems, focusing on motion effects; see Section 2 for a review.

These methods can effectively create motion effects for a *single* target class. However, using these algorithms requires caution when dealing with the scenes that involve multiple events or causes for providing motion effects of *different* classes. For example, a scene may feature apparent motions of both the camera and an object of interest, as shown in Figure 1. If these two elements move in opposite directions, adding the corresponding camera and object motion effects may cancel each other out, resulting in a weakened effect. Conversely, depending on the directions of the two causes, the added effect can become excessive. Moreover, scenes often simultaneously involve an explicit camera motion and a loud sound effect, such as an explosion. In such cases, a short and sudden motion effect synchronized with the sound effect may interfere with the camera motion effect. These observations raise a new research problem: *"How can we merge two motion effects of different classes while maintaining the desired perceptual consequences of both effects and preserving the contextual associations between the scene's content and the merged motion effect?"*

In this paper, we consider merging the motion effects of the two representative classes: camera and object. These motion effects are the most frequently used for 4D films and rides [27], as well as XR applications. We propose an algorithmic framework that merges camera and object motion effects into one motion effect while preserving the perceptual consequences of the two effects (Section 3). We validate the merging methods' performance through a user study (Section 4), which compares camera and object motion effects with their merged effects by three different weighting schemes. It is followed by a discussion on the current limitations of our work and future research topics (Section 5). This work is one of the first

*e-mail: jjpark17@postech.ac.kr
†e-mail: han0209@postech.ac.kr
‡e-mail: choism@postech.ac.kr

979-8-3503-2838-7/23/$31.00 ©2023 IEEE
DOI 10.1109/ISMAR59233.2023.00120

successful attempts to combine different kinds of motion effects for improved 4D and XR experiences.

2 Related Work

2.1 Multisensory Effects

The origins of multisensory effects can be traced back to an early experiment, Sensorama [17], which aimed to replicate real-world experiences by incorporating vision, sound, wind, scent, and motion. Participants could immerse themselves in a virtual motorcycle ride on the streets of New York. Over time, technological advancements have paved the way for more sophisticated 4D effects. Researchers have integrated environmental interactions into XR, such as temperature [9, 35, 36], motion [5, 12, 13, 19], and even taste [3, 10, 37].

In contrast to the extensive research on the effectiveness of multisensory effects, most 4D effects still need to be created manually by human experts, limiting their widespread adoption. Developing guidelines and techniques for automatically generating these effects is still in its early stages. Related research focused on 4D films, one of the most popular and widespread forms of 4D content, and games. For example, vibration effects can be generated from sounds by using psychoacoustic features [25] or from videos using a saliency map [23], or from both [29], to enhance audiovisual experiences. Appropriate 4D effects can be automatically labeled for each scene using video classification techniques [41]. This approach is useful for providing different types of haptics effects [6] and multisensory effects for 360° videos [42]. This crucial area of automatic multisensory effect authoring awaits significant research endeavors.

2.2 Motion Effects

Motion effects, as a subset of multisensory effects, enhance the perceived realism and presence in XR applications to a great extent. Initially, motion platforms were used for training pilots and astronauts. Nowadays, they are commonly adopted in entertainment applications, such as theme park rides, 4D films, and virtual reality (VR) games. Such motion effects are classified into four classes: camera, object, vibration and impulse, and context [27]. Here, a *camera motion effect* is generated by following the motion of the camera shooting the scene, and an *object motion effect* represents the motion of a character or an object of interest appearing in the scene. Camera and object motion effects are the most frequently used in 4D applications [27].

The past decade has seen expanding efforts to generate proper motion effects automatically from audiovisual content. Shin et al. [40] presented the first framework for generating camera motion effects from videos. They estimated the camera position and orientation per image and represented it as a motion effect. Lee et al. [27] introduced a more sophisticated pipeline for camera motion effects. They used the optical flow to estimate the camera trajectory more efficiently and a washout filter to make motion commands from the camera trajectory. Lim et al. [30] proposed an algorithm to improve a camera motion effect by adding a high-frequency motion representing the ground texture. Camera motion effects can also be converted to spatiotemporal vibrotactile effects on a chair [38].

As for object motion effects, Lee et al. [26] estimated the 2D trajectory of an interesting object on the screen and expressed its movement as motion effects. Goh et al. [11] proposed an automatic tagging framework that generates motion effects for each scene based on the salient movement of objects. Han et al. [16] introduced a new concept, *motion proxy*, that represents an object's 6D translations and rotations in the camera space by one 3D representative point. Object motion effects are generated using a motion cueing algorithm (MCA) with the motion proxy as input. This idea was extended to computing object motion effects for multiple articulated bodies, e.g., two fighting persons, in subsequent work [15].

Motion effects are also effective in eliciting the sensations of walking. According to Ikei et al. [18], vestibular sensations are key to providing walking sensations to seated users in XR. Amemiya et al. [1] presented methods for pseudo-walking sensations using predefined motion waveforms. Lee et al. [24] proposed a data-driven algorithm for walking sensations in different gaits. Their method collects motion data from actual people and processes the data to synthesize motion commands.

Researchers also studied the complex perceptual and emotional characteristics of motion effects. Han et al. [14] identified two primary adjective axes for motion effect perception and designed an interpolation method between motion effects to obtain a motion effect of target perceptual characteristics. Lee et al. [28] demonstrated a case for *vestibular capture*, where vestibular cues are more dominant than visual cues in motion perception. Jeong and colleagues investigated the types and strengths of the emotions elicited by motion effects in different axes and intensities under explicit consideration of the user's empathy level [20, 21].

Currently, we are aware of only one study that handles merging motion effects from different classes. Yun et al. [44] combined camera and sound motion effects to enhance the spectator experiences of watching video games. A sound motion effect is triggered when a gunfire sound is detected. It is designed to use the motion axes not predominantly used by a camera motion effect at a particular moment to avoid interference between the two effects. This handy solution benefits from sound motion effects being much shorter and stronger than camera motion effects. In contrast, camera and object motion effects have comparable lengths and intensities, and merging them without losing perceptual information is more challenging.

3 Motion Merging Algorithms

The computational procedure of our motion merging algorithms is outlined in Figure 2. Details are described in this section.

3.1 Algorithm Overview

Let $\boldsymbol{x} = (r, p, h)^T$ be an arbitrary vector in the 3-DOF motion space of *roll*, *pitch*, and *heave*, commanded to a standard 3-DOF motion platform (see Figure 1). We denote a camera motion effect by $\boldsymbol{x}^c$ and an object motion effect by $\boldsymbol{x}^o$. These two motion effects are the input to our motion merging procedure. They are assumed to have zero initial positions, as required for motion commands. Then, these two position motion effects can be equivalently expressed by their velocity vectors, $\boldsymbol{v}^c = \dot{\boldsymbol{x}}^c$ and $\boldsymbol{v}^o = \dot{\boldsymbol{x}}^o$, respectively.

We linearly combine the two input motion effects in the velocity space using time-dependent weights:

$$\boldsymbol{v}^l[n] = w^c[n]\boldsymbol{v}^c[n] + w^o[n]\boldsymbol{v}^o[n], \tag{1}$$

where $\boldsymbol{v}^l[n]$ is the linearly-combined motion velocity at the discrete time index n, and $w^c[n]$ and $w^o[n]$ are the weights of the camera and object motion effects, respectively, with $w^c[n] + w^o[n] = 1$. We perform merging in the velocity space, not position or acceleration, as velocity naturally represents motion. Note that the two weighting functions can be time-dependent. We design and evaluate three weighting schemes in this work (Section 3.2).

A motion command directly generated from $\boldsymbol{v}^l[n]$ tends to include excessive fluctuations although the sum of $w^c[n]$ and $w^o[n]$ are regulated as 1. For smoothing, we apply an additional scaling procedure that attenuates $\boldsymbol{v}^l[n]$ to $\boldsymbol{v}^a[n]$ (Section 3.3).

The final step (Section 3.4) feeds the attenuated velocity $\boldsymbol{v}^a[n]$ to a motion cueing algorithm using model predictive control (MPC). It results in a merged motion command in position, $\boldsymbol{x}^m[n]$, which respects the motion platform's limited workspace.

3.2 Weighting Methods

In principle, the optimal weights for motion effect merging in (1) hinge on the user's attention and expectation, which are determined by many factors, such as visual size, saliency, and sound. We focus

on the information contained in the camera and object motion effects, where faster changes generally signify more important movements in the scene. The weighting policies described below have different degrees of emphasizing faster changes in motion.

3.2.1 Uniform (UNI)

This uniform weighting method combines the two input motion effects in an equal ratio:

$$w^c[n] = 0.5 \quad \text{and} \quad w^o[n] = 0.5. \tag{2}$$

UNI is considered as the baseline.

3.2.2 Salient (SAL)

This policy emphasizes a more salient motion effect while regarding a motion effect with a greater velocity as more salient, such that:

$$w^c[n] = \frac{\|\boldsymbol{v}^c[n]\|}{\|\boldsymbol{v}^c[n]\| + \|\boldsymbol{v}^o[n]\|} \text{ and } w^o[n] = \frac{\|\boldsymbol{v}^o[n]\|}{\|\boldsymbol{v}^c[n]\| + \|\boldsymbol{v}^o[n]\|}. \tag{3}$$

Therefore, SAL prefers faster motion effects.

3.2.3 Selective (SEL)

The next weighting scheme selects only a dominant motion effect, which is judged by the motion velocity. For that, we first compute

$$\tilde{w}^c[n] = \begin{cases} 1 & \text{if } \|\boldsymbol{v}^c[n]\| \geq \|\boldsymbol{v}^o[n]\| \\ 0 & \text{otherwise} \end{cases},$$
$$\tilde{w}^o[n] = \begin{cases} 1 & \text{if } \|\boldsymbol{v}^c[n]\| < \|\boldsymbol{v}^o[n]\| \\ 0 & \text{otherwise} \end{cases}. \tag{4}$$

This simple binary policy requires additional processing for two reasons. First, the merged motion computed using (4) may include discontinuous values for the continuous input effects of $\boldsymbol{v}^c$ and $\boldsymbol{v}^o$. Such values can lead to a jerky motion command, potentially resulting in poor user experiences and damaging the motion platform. Second, the dominance of one motion effect must sustain for some time to impart sufficient sensory experiences.

Therefore, we apply two-step smoothing to $\tilde{w}^c[n]$ and $\tilde{w}^o[n]$ as follows. The first step aims to suppress excessive switching between the two motion effects. We process $\tilde{w}^c[n]$ and $\tilde{w}^o[n]$ so that one motion effect must prevail continuously for a predefined minimum segment length L_{min}. To this end, we segment the sequence $\tilde{w}^c[n]$ into the blocks of 0 and 1, as illustrated in Figure 3. Each block b has two neighbor blocks, b_{left} and b_{right}, of the other number. Let $L(b)$ be the length of block b. We define the relative length of b as

$$L_{re}(b) = \begin{cases} L(b) - L(b_{left}) - L(b_{right}) & \text{if } L(b) < L_{min} \\ L_{min} & \text{otherwise} \end{cases}. \tag{5}$$

The relative length $L_{re}(b)$ becomes lower when $L(b)$ is shorter but its neighbor blocks are longer. We scan the blocks to find a block b^* that has a relative length not greater than both of its neighbors.

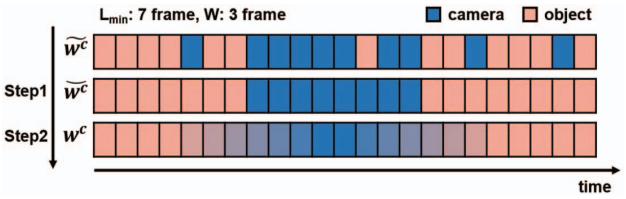


Figure 3: Smoothing process for the weighting policy SEL. For w^c, a blue cell has 1, while an orange cell has 0. A cell color between orange and blue represents an interpolated value between 1 and 0.

It indicates that b^* is relatively less significant than the adjacent blocks. We flip b^*'s elements (0→1 or 1→0) to make the three consecutive blocks, b^*_{left}, b^*, and b^*_{right}, all represent the same class of motion effect. We continue this process from the next block of b^*_{right} without updating the relative length of the merged block to prevent long blocks from absorbing subsequent blocks. When the scanning and flipping process reaches the end of the blocks, the relative lengths are recalculated. This procedure is repeated until all blocks in the sequence $\tilde{w}^c[n]$ are no shorter than L_{min}; see step 1 in Figure 3. The resulting sequence is denoted by $\check{w}^c[n]$.

The second step ensures smooth transitions at switching between $\boldsymbol{v}^c$ and $\boldsymbol{v}^o$. We copy $\check{w}^c[n]$ to $w^c[n]$ and then find a time index n^* at which $w^c[n^*] = 1$ and $w^c[n^*+1] = 0$. Then, for a window length W, we compute for $n \in [n^* - W, n^* + 1 + W]$

$$w^c[n] = \frac{1}{2}\left(\cos\left(\frac{n - (n^* - W)}{2W+1}\pi\right) + 1\right), \tag{6}$$

which smoothly interpolates $w^c[n]$ from 1 to 0 between $[n^* - W, n^* + 1 + W]$. Similarly, for the time index n^* at which $w^c[n^*] = 0$ and $w^c[n^*+1] = 1$, we interpolates $w^c[n]$ from 0 to 1 by

$$w^c[n] = \frac{1}{2}\left(-\cos\left(\frac{n - (n^* - W)}{2W+1}\pi\right) + 1\right), \tag{7}$$

for $n \in [n^* - W, n^* + 1 + W]$. The other elements in $w^c[n]$ are left intact. This smoothing process is visualized in step 2 of Figure 3.

Finally, $w^o[n]$ is computed by

$$w^o[n] = 1 - w^c[n]. \tag{8}$$

This pair of $w^c[n]$ and $w^o[n]$ is denoted by SEL (selective).

3.3 Motion Attenuation

A motion command computed directly from the linearly combined velocity $\boldsymbol{v}^l$ in (1) tends to exhibit frequent and excessive motion effects regardless of the weighting policy used. This problem is inevitable because the two source motion effects are triggered by different events in the scene. We provide a means to control the instability by an additional step of motion attenuation, as follows:

$$\boldsymbol{v}^a = \|\boldsymbol{v}^l\|^\gamma \frac{\boldsymbol{v}^l}{\|\boldsymbol{v}^l\|}, \tag{9}$$

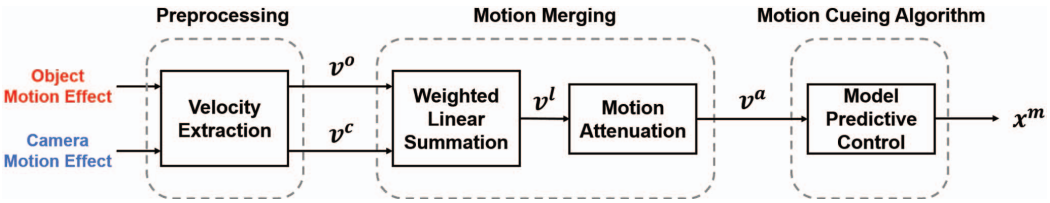


Figure 2: Computational procedure of motion effect merging.

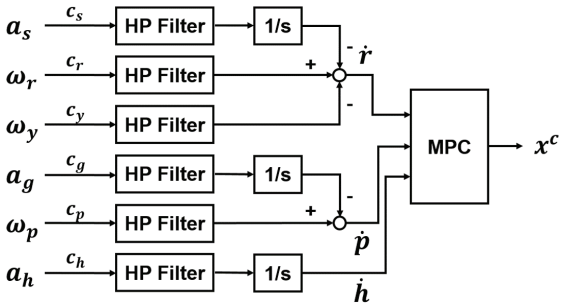


Figure 4: Motion cueing algorithm used to generate camera motion effects from camera motion ($\boldsymbol{a}$ and $\boldsymbol{\omega}$). This MCA is improved from the original one in Lee et al. [27, Figure 7]. HP Filter: High-pass filter.

where γ is the motion attenuation factor between 0 and 1.

3.4 Position Command Computation

Finally, we compute the motion command in position, $\boldsymbol{m}[n]$, to the motion platform. To this end, the attenuated velocity of the added motion, $\boldsymbol{v}^a$ in (9), is processed by a standard motion cueing algorithm using MPC[1]. We use an MPC formulation with velocity input [16]:

$$\boldsymbol{x}^m = \underset{\boldsymbol{m}}{\text{argmin}} \quad \|\boldsymbol{v}^a - \dot{\boldsymbol{m}}\|^2 + \varepsilon\|\boldsymbol{m}\|^2 \qquad (10)$$
$$\text{subject to} \quad |\boldsymbol{m}| \leq \boldsymbol{x}_{max},$$

where $\boldsymbol{x}^m$ is the final motion command merging the two input motion effects, $\boldsymbol{x}^c$ and $\boldsymbol{x}^o$. ε is a coefficient that controls the motion platform's returning speed to the neutral position. $\boldsymbol{x}_{max} = (r_{max}, p_{max}, h_{max})^T$ is the maximum range of the motion platform. The inequality in the constraint in (10) represents element-wise operations. The merged motion effect $\boldsymbol{x}^m$ found by the MPC optimization in (10) minimizes the velocity difference between the desired motion (represented by $\boldsymbol{v}^a$) and the output motion $\dot{\boldsymbol{m}}$ while respecting the constrained range of the motion platform. $\boldsymbol{x}^m$ is the final command to the motion platform.

3.5 Implementation Details

For implementation and evaluation, we use a commercial 3-DOF (degree of freedom) motion chair (CJ 4DPlex, NX1; see Figure 7) used in theaters. The chair is for four people. It has an operating range of roll $\pm 4°$, pitch $\pm 7°$, and heave ± 4 cm.

We use the algorithm by Lee et al. [27] to generate camera motion effects from a video. This method estimates the camera's 3D linear accelerations and angular velocities from the image sequences. Then, it synthesizes a motion effect by feeding the camera motion variables to a washout filter (cutoff frequencies 2.5 Hz and 1 Hz for linear accelerations and angular velocities, respectively). In our implementation, the image-based camera motion estimation is replaced by the ground-truth camera velocity (available from Unity 3D; see Section 4.1.2). The original "limit and enhance" operations are also improved by using an MPC module (the same one as in (10)), as shown in Figure 4. These motion effects are referred to as CAM.

Object motion effects are made using Han et al.'s algorithm [16]. This method expresses the object's movement and rotation by a single motion proxy and synthesizes a motion effect by feeding the motion proxy to MPC. We use the MPC with velocity input [16, Eq. (10)], which resulted in the best user experiences in their user study. These motion effects are denoted by OBJ.

[1]MPC is a control methodology that optimizes the current control input based on a process model and a future trajectory while satisfying constraints [4]. MPC attempts to minimize the difference between the reference output and the output produced by the process model.

As described earlier, our motion merging algorithms have four parameters: L_{min}, W, γ, and ε. Their values found by pilot tests for our motion chair are shown in Table 1.

Table 1: Parameter values of the motion merge algorithms used for all experiments and user study.

L_{min}	W	γ	ε
31 frames*	15 frames	0.55	0.3

* The videos are processed at 30 frame/s.

3.6 Examples

Figure 5 shows motion command examples produced by the three weighting methods. The motion commands are for the roll axis (left-right rotation on the screen). The two input motion commands, CAM and OBJ, are shown in blue and red dashed lines, respectively. The three output motion commands of motion merging, UNI, SAL, and SEL, are shown in thick purple lines.

Between 12 s and 15 s, OBJ shows a large left-to-right motion (from negative to positive), while CAM is mostly constant. Thus, the object velocity overwhelms the camera velocity in this interval. In response, all three merged motion effects track OBJ, albeit at varying degrees. UNI yields the smallest movement, approximately half OBJ, whereas SAL and SEL result in much greater motions. Around 19 s and 33 s, CAM and OBJ vary in different directions. In these regions, UNI and SAL lead to relatively small motions due to the cancellation between CAM and OBJ. However, SEL demonstrates a substantially large motion aligned with the dominant motion. In general, the intensity of generated motion is the lowest with UNI, followed by SAL, and the greatest with SEL.

4 User Study

We conducted a user study to compare motion effects produced by single-class authoring algorithms and their counterparts merged using our algorithms. This experiment was approved by the Institutional Review Board at the author's institution (PIRB-2022-E019).

4.1 Methods

4.1.1 Participants

Twenty volunteers (6 females and 14 males) with normal sensory and motor ability participated in the user study. Their average age was 22.8 ± 3.3 years. Before the experiment, participants were provided

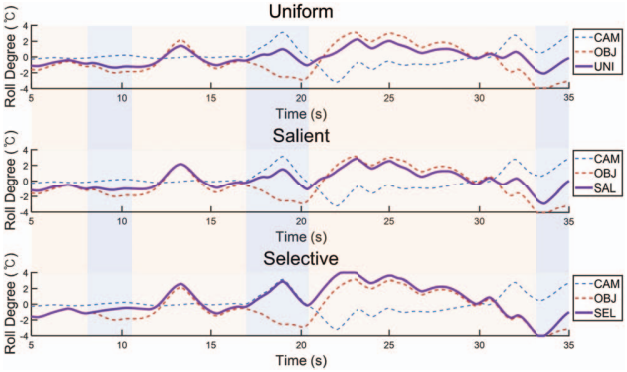
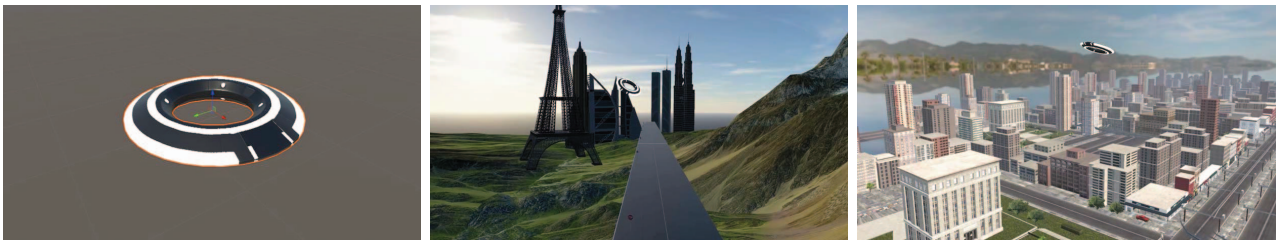


Figure 5: Motion command examples (roll) of the three weighting algorithms. The background is colored blue where CAM is dominant and red where OBJ is dominant.



(a) Object (b) Airstrip (c) City Day

(d) City Night (e) Lake (f) Valley

Figure 6: Moving object (a) and the backgrounds of five videos (b-f).

Table 2: Summary of the camera and object motions appearing in the videos.

#	Animation	Description	Duration (s)
1	Airstrip	The camera takes off, moves up and down smoothly, and then rotates the buildings dynamically. The object leads the camera at a close distance while moving relatively randomly.	38.1
2	City Day	The camera takes off and then rotates 360 degrees in diagonal directions like a roller coaster. The object moves slowly and stays mostly in view, leaving and entering the scene only once.	35.4
3	City Night	The camera takes off and swings left and right quite violently like a pendulum ride. The object moves relatively slowly and repeats leaving and entering the scene as the camera moves faster.	36.2
4	Lake	The camera takes off and then rotates around the lake counterclockwise. The object moves closely in front of the camera while leaving and entering the scene twice.	35.1
5	Valley	The camera falls to the valley, levels off with little vertical motion, and then rises up along a waterfall. The object appears after the camera's plunge and then repeatedly moves left and right.	27.6

with a written document detailing the objective and procedure of the study. Following this, they signed a consent form to confirm their participation agreement. Each participant received a compensation of KRW 30,000 (≈ USD 23).

4.1.2 Experimental Conditions

We made five videos using a 6-DOF flight simulator available in Unity 3D. Each video included one flying disk-like object (Figure 6a). Both the camera and object motions were smooth, obeying the laws of physics provided by the simulator's physics engine. The five videos are available in the supplemental material. Their screenshots are presented in Figure 6b–f. The behaviors of the camera and object were designed to cover a diverse range of motion, as described in Table 2. The object often enters or leaves the camera's view in the videos, as in natural scenes. During such cases, we set the object motion velocity to zero, which keeps the object motion command constant in position. Each video's duration was made at approximately 30 s. The five videos also varied considerably in visual factors, such as landscape and lighting, as they may affect the user's visual perception of the camera and object motion.

We prepared five motion effects, CAM, OBJ, UNI, SAL, and SEL, for each of the five videos. The former two were single-class motion effects, and the latter three were the motion effects merged from the two single-class effects. As a result, this user study had 25 experimental conditions. The camera and object velocities used in authoring the motion effects were directly extracted from the motion trajectories of the camera and object in Unity 3D.

4.1.3 Procedure

A participant sat on the second seat from the left of the four-person motion platform (Figure 7). Participants wore noise-canceling headphones (Bose, QC 45) playing white noise to block out the platform's operational sounds. A video was projected onto a 94-inch screen using one channel of a polarized projector (Epson, EB-W16SK).

The experiment consisted of training and main sessions. The training session aimed to familiarize participants with the experiment's stimuli by exposing them to one video (#3) and its five motion effects. Video 3 was chosen as it was apt to present individual camera and object motion effects more clearly. The main session comprised five blocks, each corresponding to one video. Their order was randomized per participant. Each block had two repetitions of five trials each. In each trial, participants perceived the video and one motion effect and then evaluated the combination's user experiences by responding to a questionnaire (Section 4.1.4) for 30 s. The order of the five motion effects in both repetitions was balanced across participants using a Latin square. Participants recorded rough answers to the questionnaire during the first repetition of trials and then finalized them during the second repetition. Participants were

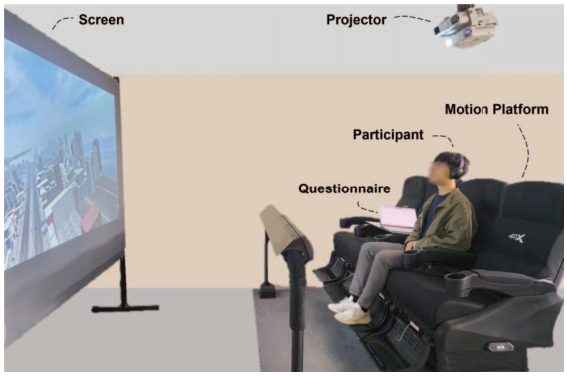


Figure 7: Experimental setup.

given a 2-min break after each block. The entire experiment took approximately 80 min.

4.1.4 Questionnaire

Each motion effect was evaluated using a questionnaire consisting of the following seven questions: Q1. *Harmony*: The motion effect matched the video; Q2. *Camera Expressiveness*: The motion effect described the important parts of the camera's movements; Q3. *Object Expressiveness*: The motion effect described the important parts of the object's movements; Q4. *Discomfort*: The motion effect made me feel uncomfortable (dizziness, nausea, or fatigue); Q5. *Distraction*: The motion effect interfered with focusing on appreciating the video; Q6. Preference rank: Rank the five motion effects in the order of preference; and Q7. Free comments: Provide any comments about the motion effects, video, and experiment.

Q1–Q5 were answered in arbitrary real numbers from 0 to 100. The score meanings were labeled as: 0–strongly disagree, 25–disagree, 50–neutral, 75–agree, and 100–strongly agree. Q7 was a free essay question. Participants responded to Q1–Q5 and Q7 after perceiving each combination of video and motion effect. For Q6, participants ranked the five motion effects in descending order of preference without ties, per video at the end of each block.

4.1.5 Data analysis

We computed another measure, *Scene Expressiveness*, as the average of the scores of Q2 (*Camera Expressiveness*) and Q3 (*Object Expressiveness*). The scores of the two negative questions Q4 (*Discomfort*) and Q5 (*Distraction*) were inverted for the notation consistency between a high score and positive user experience. Those inverted measures are denoted by a ¬ sign, e.g., $\neg Discomfort$ and $\neg Distraction$. Also, the rank data obtained from Q6 (Preference) was inverted to a preference score by $(6 - rank)$.

4.2 Results

4.2.1 User Experience Measures

We conducted a two-way repeated-measures ANOVA with two independent factors of *Motion Effect* and *Video* on each of the collected scores of *Harmony*, *Camera Expressiveness*, *Object Expressiveness*, *Scene Expressiveness*, ¬*Discomfort*, and ¬*Distraction*. The results are summarized in Table 3. *Motion Effect* had statistically significant effects on all six measures. *Video* was significant only for *Harmony* and *Camera Expressiveness*. Their interaction term was significant for all measures except *Object Expressiveness*.[2] Figure 8 shows the mean scores of the six measures averaged over the five videos for the five motion effects. Figure 9 shows the mean scores of *Harmony* and *Camera Expressiveness*, for which *Video* was significant, of the five motion effects for each of the five videos. We also ran the SNK tests for post-hoc multiple comparisons on the significant main effects and represented the results in the plots. These results are also summarized in Table 4.

Among the five kinds of motion effects tested, SAL and SEL led to the best overall scores in all six user experience measures, while OBJ resulted in the lowest scores. For the five videos, only Video 5 showed significantly lower scores in only *Harmony* and *Camera Expressiveness*. Thus, the effect of *Video* was generally weak.

4.2.2 Preference

The scores of *Preference* are ordinal data. Hence, for statistical analysis, we only tested the main factor of *Motion Effect* over all videos and for each video using the Friedman test. It was followed by multiple-comparison post-hoc tests using the procedure outlined in [39]. In the results, *Motion Effect* was significant for *Preference* in all cases: $\chi^2(4) = 58.16$, $p < 0.0001$ for all videos, $\chi^2(4) = 23.08$, $p < 0.0001$ for Video 1, $\chi^2(4) = 10.64$, $p = 0.031$ for Video 2, $\chi^2(4) = 23.24$, $p < 0.0001$ for Video 3, $\chi^2(4) = 20.24$, $p = 0.0004$ for Video 4, and $\chi^2(4) = 12.32$, $p = 0.0151$ for Video 5.

Figure 10 shows the preference scores. Generally, *Preference* was the highest with SAL and SEL and the lowest with OBJ. For example, the order of *Preference* over all videos was ${}_A\{\mathsf{SEL} > {}_B\{\mathsf{SAL}\}_A > {}_C\{\mathsf{CAM}\}_B > {}_D\{\mathsf{UNI}\}_C > \mathsf{OBJ}\}_D$.

4.3 Discussion

4.3.1 Comparisons Between Motion Effects

The two merged motion effects, SAL and SEL, consistently received the highest scores in almost all evaluation metrics. Both methods prioritize the dominant motion effect during merging. Comments from the participants highlighted that these motion effects provided individual camera and object motion effects well, as well as their combinations. For example, P10 (Participant 10) said, "The object's motion in front of me was more noticeable than the background (camera) motion. The camera motion and the object motion were well combined in the effects. The effects were dramatic because they matched the camera's view shift." P2 said, "This effect was good because it moved in the direction I wanted it to move, and I could feel the acceleration (from the camera motion)," complimenting the seamlessly integrated camera and object motion effects. It is noted that SEL demonstrated greater scores in almost all measures than SAL, including *Preference*, but with no significant differences.

[2]Further analysis of the significant interaction terms resulted in excessively complicated interpretations. They are not reported here for brevity, also because our main interest is in *Motion Effect*.

Table 3: Results of two-way ANOVA.

Factor	*dM*	*dE*	*Harmony*		*Cam. Exp.*		*Obj. Exp.*		*Scene Exp.*		*Discomfort*		*Distraction*	
			F	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*
E	4	76	14.85	**<.0001**	15.97	**<.0001**	3.82	**0.0070**	13.41	**<.0001**	5.68	**0.0005**	10.53	**<.0001**
V	4	76	5.70	**0.0005**	3.86	**0.0066**	0.33	0.8603	1.67	0.1646	0.85	0.4964	1.95	0.1109
$E \times V$	16	304	3.75	**<.0001**	2.82	**0.0003**	1.30	0.1972	2.56	**0.0010**	1.91	**0.0194**	2.98	**0.0001**

* *E*: *Motion Effect*, *V*: *Video*, *dM* and *dE*: Degree of freedoms for the model and the error. Statistically significant cases are marked with bold face at *p* values.

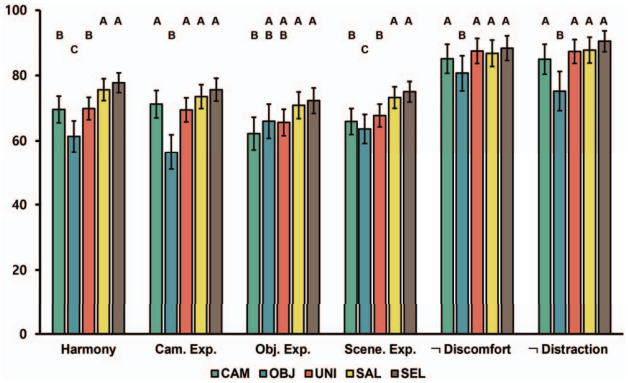


Figure 8: User experience measures for the five motion effects averaged over all five videos. Means not grouped by the same letter are significantly different from each other by the SNK tests.

Table 4: Order of the conditions in the significant factors for each user experience measure.

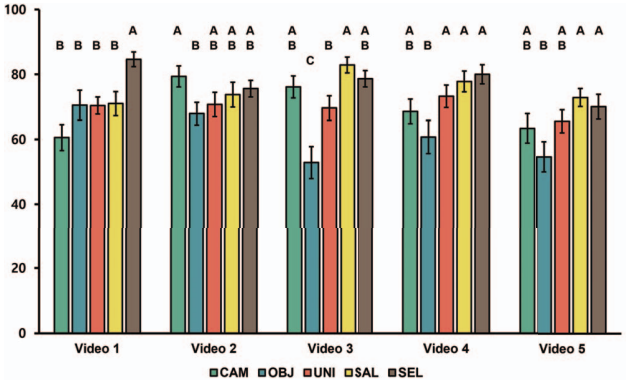
Measure	Ft.	Order of Conditions
Harmony	*E*	${}_A\{\mathsf{SEL} > \mathsf{SAL}\}_A > {}_B\{\mathsf{UNI} > \mathsf{CAM}\}_B > {}_C\{\mathsf{OBJ}\}_C$
	V	${}_A\{\#1 > \#2 > \#3 > \#4\}_A > {}_B\{\#5\}_B$
Cam.	*E*	${}_A\{\mathsf{SEL} > \mathsf{SAL} > \mathsf{CAM} > \mathsf{UNI}\}_A > {}_B\{\mathsf{OBJ}\}_B$
Exp.	*V*	${}_A\{\#1 > \#2 > \#3 > \#4\}_A > {}_B\{\#5\}_B$
Obj. Exp.	*E*	${}_A\{\mathsf{SEL} > \mathsf{SAL} > {}_B\{\mathsf{OBJ} > \mathsf{UNI}\}_A > \mathsf{CAM}\}_B$
Scn. Exp.	*E*	${}_A\{\mathsf{SEL} > \mathsf{SAL}\}_A > {}_B\{\mathsf{UNI} > \mathsf{CAM}\}_B > {}_C\{\mathsf{OBJ}\}_C$
¬Discom.	*E*	${}_A\{\mathsf{SEL} > \mathsf{UNI} > \mathsf{SAL} > \mathsf{CAM}\}_A > {}_B\{\mathsf{OBJ}\}_B$
¬Distrac.	*E*	${}_A\{\mathsf{SEL} > \mathsf{SAL} > \mathsf{UNI} > \mathsf{CAM}\}_A > {}_B\{\mathsf{OBJ}\}_B$

- Ft: Factor, *E*: *Motion Effect*, *V*: *Video*.
- The conditions enclosed by ${}_X\{\cdots\}_X$ do not have statistically significant differences from each other.
- $x \in X = {}_X\{\cdots\}_X$ and $y \in Y = {}_Y\{\cdots\}_Y$ have a significant difference if $x \notin Y$ and $y \notin X$.

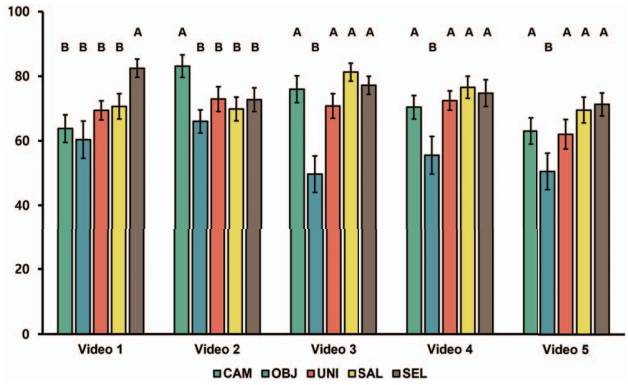
The merging effectiveness of SAL and SEL can also be confirmed by the results that they obtained higher scores in *Camera Expressiveness* than CAM and higher scores in *Object Expressiveness* than OBJ (but without statistical significance). These observations suggest that the features delivering the sensations of camera and object motions are adequately preserved in the merged effects, SAL and SEL. Consequently, SAL and SEL attained significantly higher *Scene Expressiveness* than the other three motion effects (Figure 8).

The other type of merged motion effect, UNI, did not receive as good scores as SAL or SEL. Ps pointed out fewer movements in UNI as a major problem. P1 said, "I felt fewer movements than SEL and SAL." and P18 said, "It seems like there was less motion overall." This issue can be attributed to the weaker motion intensity of UNI. When only one motion effect is prominent and the other effect is close to zero, UNI has an amplitude similar to half of the dominant effect's amplitude. It can reduce the general impression of motion intensity and frequency that users accumulate over time.

Between the two single-class motion effects, CAM generally garnered better user experience scores than OBJ. It may be because users considered themselves as passengers in the plane and regarded camera motion effects as more indispensable than object motion effects, or the result originated from the characteristics of the specific



(a) Harmony



(b) Camera Expressiveness

Figure 9: Scores of Harmony (a) and Camera Expressiveness (b) for each video. *Video* was significant for these two measures. Means not grouped by the same letter are significantly different by the SNK tests.

scenes used in the user study. Hence, based on the present results, we do not draw any conclusions about the relative importance between camera and object motion effects.

4.3.2 Video-Dependent Results

Video 5 resulted in significantly lower scores in *Harmony* and *Camera Expressiveness* than the other four videos (Figure 9). We found that the camera motion in Video 5 was smoother than in the other videos; see the supplemental video. Its camera motion effects varied under half of the motion platform's range in all axes. Although Video 5 features an initial deep fall, the motion platform's heave range does not seem sufficient to express the fall.[3] Many Ps mentioned awkwardness and diminished immersion due to the relative lack of motion effects in Video 5.

In Video 1, OBJ was preferred to CAM (Figure 10) unlike the other videos. Many Ps mentioned that they focused on the flying object in this video. Video 1 has the least background landmarks and features from which Ps can visually estimate the camera motion (Figure 6 and the supplemental video). It may have diverted Ps' attention more to the object's movement.

In Video 2, CAM tended to show better scores than the merged effects SAL and SEL (Figure 9 and 10). Video 2 includes a distinctive camera motion of 360° rotations in the air (Figure 11 and the supplemental video). P2 said, "I think this effect was the most realistic for the scene where the camera turned a full circle." P16 also said, "The 360-degree turn feels smooth." It also appears that an object motion

[3]Our motion platform is for four persons in theaters, and such motion chairs typically have a small range of motion in heave (up-down direction).

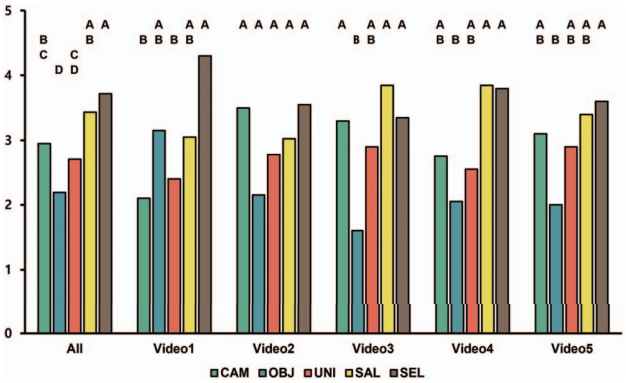


Figure 10: Average preference scores for the five motion effects over all videos (left) and for each video (right five). Means not grouped by the same letter are significantly different.

effect added to such a prominent camera motion effect can adversely affect the quality of the merged motion effect; for example, compare CAM and SEL with SAL for Video 2 in Figure 10.

4.3.3 Discomfort and Distraction

Discomfort is a negative factor affecting motion sickness, and motion effects with higher magnitudes are more likely to induce motion sickness [22]. However, such a tendency was not observed in our results. SAL and SEL, which exhibit faster and stronger motion commands than other effects, received slightly better scores for discomfort than the single-class motion effects (Figure 8). Conversely, OBJ obtained a worse score in discomfort.

To discover reasons for these intriguing results, we first noted that motion stimuli with high energy in a very low-frequency band around 0.2 Hz are critical for motion sickness [34]. However, our motion stimuli did not have significant differences in the low-frequency intensities in their power spectra. Another potential reason is that the discordance between visual and vestibular stimuli frequently causes motion sickness. The merged effects, SAL and SEL, provided vestibular stimuli more harmonious with visual stimuli than the single-class effects. This advantage could have contributed to the more comfortable sensory experiences afforded by SAL and SEL. Additionally, our motion effects were not excessively intense, with a maximum tilting of 4°.

As for distraction, the merged motion effects, SAL and SEL, also led to better scores than the single-class effects. This result may be accounted for by the merged effects' higher harmony (perceived match with the video) scores, involving less distraction.

5 Limitations and Future Work

We tried to encompass a diverse set of situations in the experimental evaluation. However, we still need to test our algorithms for more natural and various scenes to confirm the ecological validity. For instance, visual scenes in first-person shooter (FPS) games entail considerably more vigorous movements, including actions like rolling forward, executing knife attacks, and firing projectiles, than the movements simulated in our experiment. Hence, testing in a broader range of scenarios with varying complexity would be worthwhile to refine our algorithms. To this end, we have been testing and tuning the parameters of our merging algorithm for the regular film scenes that exhibit clear camera and object motions. Example results are appended in the supplemental video.

The participants of the user study were all young in their 20s and biased to males with no motion disorders. Age, gender, and motion disorder are important factors for the robustness and generalizability of our algorithms and findings, all requiring further study.

Figure 11: Image sequence of 360° flip in video 2.

Our merging algorithms use motion velocity to assess the motion's importance. Although effective for their purpose, they are independent of other visual information available in the scene, such as visual saliency, scene flow, and object identity. Including such high-level information is expected to better estimate the user's mental model in watching 4D films and provide additional insights into motion effect merging. For that, the visual display used can also matter significantly. All of these can contribute to developing more effective, comprehensive, and robust motion merging algorithms.

Lastly, as the final goal of our research, it would be beneficial to incorporate other classes of motion effects to understand their interactions and derive an ultimate merging scheme.

6 Conclusions

This paper reports an initial attempt to develop effective algorithms for merging two kinds of motion effects: camera and object. Our algorithms combine camera and object motion velocities by a linear weighted sum. Three weighting policies are designed and tested, which have different extents of emphasizing faster motions. A user study assessing the effectiveness of the three methods advocates that a faster motion effect deserves a higher weight in motion effect merging. Such merged motion effects are consistently evaluated as more harmonious and expressive while causing less discomfort and distraction than single-class motion effects, thereby leading to a higher preference of users. Our algorithms are directly applicable to 4D films and rides, as well as other motion-enabled XR applications. In future work, we plan to extend our algorithms to encompass various factors for devices and participants, incorporate the scene's context better, and examine the users' perceptual and emotional responses in more general scenarios.

Acknowledgments

This work was supported by the Samsung Research Funding & Incubation Center (SRFC-IT1802-05), also by the Culture Technology R&D Program (CTRD R2021040136) from the Korea Creative Content Agency and the Mid-Career Researcher Program (2022R1A2C2091161) from the National Research Foundation.

References

[1] T. Amemiya, M. Kitazaki, and Y. Ikei. Pseudo-sensation of walking generated by passive whole-body motions in heave and yaw directions. *IEEE Transactions on Haptics*, 13(1):80–86, 2020. doi: 10.1109/TOH.2020.2965937

[2] Blooloop. Mediamation expands MX4D on five continents, 2019. [Online; accessed 10-June-2023].

[3] J. Brooks, P. Lopes, J. Amores, E. Maggioni, H. Matsukura, M. Obrist, R. Lalintha Peiris, and N. Ranasinghe. Smell, taste, and temperature interfaces. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. ACM, 2021. Article 76, 6 pages. doi: 10.1145/3411763.3441317

[4] E. F. Camacho and C. B. Alba. *Model predictive control*. Springer Science & Business Media, 2013.

[5] L.-P. Cheng, P. Lühne, P. Lopes, C. Sterz, and P. Baudisch. Haptic turk: A motion platform based on people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 3463–3472. ACM, 2014. doi: 10.1145/2556288.2557101

[6] C.-H. Chou, Y.-S. Su, C.-J. Hsu, K.-C. Lee, and P.-H. Han. Design of desktop audiovisual entertainment system with deep learning and

haptic sensations. *Symmetry*, 12(10), 2020. Article 1718, 14 pages. doi: 10.3390/sym12101718

[7] CJ. CGV 4DX surpasses 700 locations worldwide, 2019. [Online; accessed 10-June-2023].

[8] W. contributors. D-Box technologies, 2023. [Online; accessed 10-June-2023].

[9] A. C. da Silveira, E. C. Rodrigues, E. B. Saleme, A. Covaci, G. Ghinea, and C. A. Santos. Thermal and wind devices for multisensory human-computer interaction: An overview. *Multimedia Tools and Applications*, pp. 1–28, 2023. doi: 10.1007/s11042-023-14672-y

[10] A. S. Duggal, R. Singh, A. Gehlot, M. Rashid, S. S. Alshamrani, and A. S. AlGhamdi. Digital taste in mulsemedia augmented reality: Perspective on developments and challenges. *Electronics*, 11(9), 2022. Article 1315, 13 pages. doi: 10.3390/electronics11091315

[11] E. Goh, D. Kim, S. Oh, and C.-B. Sohn. Automatic effect generation method for 4D films. *International Journal of Computing and Digital Systems*, 9(2):281–288, 2020. doi: 10.12785/ijcds/090213

[12] J. Gugenheimer, D. Wolf, E. R. Eiriksson, P. Maes, and E. Rukzio. GyroVR: Simulating inertia in virtual reality using head worn flywheels. In *Proceedings of the Symposium on User Interface Software and Technology*, p. 227–232. ACM, 2016. doi: 10.1145/2984511.2984535

[13] J. Gugenheimer, D. Wolf, G. Haas, S. Krebs, and E. Rukzio. SwiVR-Chair: A motorized swivel chair to nudge users' orientation for 360 degree storytelling in virtual reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, p. 1996–2000. ACM, 2016. doi: 10.1145/2858036.2858040

[14] S. Han, J. Lee, G. Yun, S. H. Han, and S. Choi. Motion effects: Perceptual space and synthesis for specific perceptual properties. *IEEE Transactions on Haptics*, 15(3):626–637, 2022. doi: 10.1109/TOH.2022.3196950

[15] S. Han, J. Park, and S. Choi. Generating haptic motion effects for multiple articulated bodies for improved 4D experiences: A camera space approach. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 2023. Article 81, 17 pages. doi: 10.1145/3544548.3580727

[16] S. Han, G. Yun, and S. Choi. Camera space synthesis of motion effects emphasizing a moving object in 4D films. In *Proceedings of the IEEE Virtual Reality and 3D User Interfaces*, pp. 670–678. IEEE, 2021. doi: 10.1109/VR50410.2021.00093

[17] M. L. Heilig. Sensorama simulator. *US PAT. 3,050,870*, 1962.

[18] Y. Ikei, S. Kato, K. Komase, S. Imao, S. Sakurai, T. Amemiya, M. Kitazaki, and K. Hirota. Vestibulohaptic passive stimulation for a walking sensation. In *Proceedings of the IEEE Virtual Reality Conference*, pp. 185–186. IEEE, 2016. doi: 10.1109/VR.2016.7504715

[19] D. Jain, M. Sra, J. Guo, R. Marques, R. Wu, J. Chiu, and C. Schmandt. Immersive terrestrial scuba diving using virtual reality. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, p. 1563–1569. ACM, 2016. doi: 10.1145/2851581.2892503

[20] D. Jeong, S. H. Han, D. Y. Jeong, K. Kwon, and S. Choi. Investigating 4D movie audiences' emotional responses to motion effects and empathy. *Computers in Human Behavior*, 121:106797, 2021. doi: 10.1016/j.chb.2021.106797

[21] D. Y. Jeong, S. H. Han, S. Choi, D. Jeong, and K. Kwon. Investigating perceived emotions and affects of a scene, and the user satisfaction with motion effects in 4D movies. *International Journal of Industrial Ergonomics*, 85:103173, 2021. doi: 10.1016/j.ergon.2021.103173

[22] J. A. Joseph and M. J. Griffin. Motion sickness: Effect of the magnitude of roll and pitch oscillation. *Aviation, Space, and Environmental Medicine*, 79(4):390–396, 2008. doi: 10.3357/ASEM.2196.2008

[23] M. Kim, S. Lee, and S. Choi. Saliency-driven real-time video-to-tactile translation. *IEEE Transactions on Haptics*, 7(3):394–404, 2014. doi: 10.1109/TOH.2013.58

[24] H. Lee, S. Oh, and S. Choi. Data-driven rendering of motion effects for walking sensations in different gaits. *IEEE Transactions on Haptics*, 15(3):547–559, 2022. doi: 10.1109/TOH.2022.3176964

[25] J. Lee and S. Choi. Real-time perception-level translation from audio signals to vibrotactile effects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 2567–2576. ACM, 2013. doi: 10.1145/2470654.2481354

[26] J. Lee, B. Han, and S. Choi. Interactive motion effects design for a moving object in 4D films. In *Proceedings of the ACM Conference on Virtual Reality Software and Technology*, p. 219–228. ACM, 2016. doi: 10.1145/2993369.2993389

[27] J. Lee, B. Han, and S. Choi. Motion effects synthesis for 4D films. *IEEE Transactions on Visualization and Computer Graphics*, 22(10):2300–2314, 2016. doi: 10.1109/TVCG.2015.2507591

[28] J. Lee, S. H. Han, and S. Choi. Sensory cue integration of visual and vestibular stimuli: a case study for 4D rides. *Virtual Reality*, pp. 1–13, 2023. doi: 10.1007/s10055-023-00762-7

[29] Y. Li, Y. Yoo, A. Weill-Duflos, and J. Cooperstock. Towards context-aware automatic haptic effect generation for home theatre environments. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. ACM, 2021. Article 13, 11 pages. doi: 10.1145/3489849.3489887

[30] B. Lim, S. Han, and S. Choi. Image-based texture styling for motion effect rendering. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. ACM, 2021. Article 20, 10 pages. doi: 10.1145/3489849.3489854

[31] MaxFlight. Maxflight, 2020. [Online; accessed 10-June-2023].

[32] Motionsystems. Motion systems - motion platforms for simulators, 2023. [Online; accessed 10-June-2023].

[33] Nextlevelracing. Next level racing — advanced simulation products, 2023. [Online; accessed 10-June-2023].

[34] J. F. O'Hanlon and M. E. McCauley. Motion sickness incidence as a function of the frequency and acceleration of vertical sinusoidal motion. Technical report, Canon Research Group Inc. Human Factors Research Division, 1973.

[35] R. L. Peiris, W. Peng, Z. Chen, L. Chan, and K. Minamizawa. ThermoVR: Exploring integrated thermal haptic feedback with head mounted displays. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, p. 5452–5456. ACM, 2017. doi: 10.1145/3025453.3025824

[36] N. Ranasinghe, P. Jain, N. Thi Ngoc Tram, K. C. R. Koh, D. Tolley, S. Karwita, L. Lien-Ya, Y. Liangkun, K. Shamaiah, C. Eason Wai Tung, C. C. Yen, and E. Y.-L. Do. Season traveller: Multisensory narration for enhancing the virtual reality experience. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, p. 1–13. ACM, 2018. doi: 10.1145/3173574.3174151

[37] N. Ranasinghe, R. Nakatsu, H. Nii, and P. Gopalakrishnakone. Tongue mounted interface for digitally actuating the sense of taste. In *Proceedings of International Symposium on Wearable Computers*, pp. 80–87. IEEE, 2012. doi: 10.1109/ISWC.2012.16

[38] J. Seo, S. Mun, J. Lee, and S. Choi. Substituting motion effects with vibrotactile effects for 4D experiences. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–6. ACM, 2018. Paper 428. doi: 10.1145/3173574.3174002

[39] M. R. Sheldon, M. J. Fillyaw, and W. D. Thompson. The use and interpretation of the friedman test in the analysis of ordinal-scale data in repeated measures designs. *Physiotherapy Research International*, 1(4):221–228, 1996. doi: 10.1002/pri.66

[40] S. Shin, B. Yoo, and S. Han. A framework for automatic creation of motion effects from theatrical motion pictures. *Multimedia Systems*, 20:327–346, 2014. doi: 10.1007/s00530-013-0322-4

[41] T. S. Siadari, M. Han, and H. Yoon. 4D effect video classification with shot-aware frame selection and deep neural networks. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 1148–1155. IEEE, 2017.

[42] P. Szabo, A. Simiscuka, S. Masneri, M. Zorrilla, and G.-M. Muntean. A CNN-based framework for enhancing 360 VR experiences with multisensorial effects. *IEEE Transactions on Multimedia*, pp. 1–14, 2022. doi: 10.1109/TMM.2022.3157556

[43] Z. Yuan, S. Chen, G. Ghinea, and G.-M. Muntean. User quality of experience of mulsemedia applications. *ACM Transactions on Multimedia Computing, Communications, and Application*, 11(1s), 2014. Article 15, 19 pages. doi: 10.1145/2661329

[44] G. Yun, H. Lee, S. Han, and S. Choi. Improving viewing experiences of first-person shooter gameplays with automatically-generated motion effects. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 2021. Article 320, 14 pages. doi: 10.1145/3411764.3445358