*Citation for published version:*
Kerr, W, Jicol, C, Haines, TSF & Li, W 2024, Camera Chameleon - The Creative Impact of Tracked Tangible Interfaces for Virtual Film Pre-Production. in *2024 IEEE International Conference on Multimedia and Expo, ICME 2024.* Proceedings - IEEE International Conference on Multimedia and Expo, IEEE, U. S. A., pp. 1-6, 2024 IEEE International Conference on Multimedia and Expo, ICME 2024, Niagra Falls, Canada, 15/07/24. https://doi.org/10.1109/ICME57554.2024.10687756

*DOI:*
10.1109/ICME57554.2024.10687756

*Publication date:*
2024

*Document Version*
Peer reviewed version

# Camera Chameleon — The Creative Impact of Tracked Tangible Interfaces for Virtual Film Pre-Production

Will Kerr, Crescent Jicol, Tom S. F. Haines, Wenbin Li
*Department of Computer Science*
*University of Bath*
Bath, UK
{wk307, cj406, tsfh20, wl281}@bath.ac.uk

***Abstract*—Shows such as The Mandalorian have used Virtual Production (VP) to great creative advantage. This combination, of tracked cameras and virtual environments, gives film crews the flexibility to rapidly iterate ideas and discover the best telling of their story within the time constraints of movie production. We explore if the same ideas can translate to pre-production, specifically if similar advantages can be found during the storyboarding phase. Developing a novel virtual production interface named Camera Chameleon, we compare user interaction against traditional keyboard/mouse peripherals when creating storyboards. Similar expert-rated quality is measured on storyboard output, however, users report increased enjoyment and complete shots more quickly. Novice users explore more when using our approach, demonstrating its creative potential.**

***Index Terms*—Virtual Cinematography, Cinematography, Motion Tracking, Human-Computer Interaction**

## I. INTRODUCTION

Within virtual cinematography keyboard/mouse (KM) remains the default tool, though many virtual environments do permit other hardware inputs, e.g. Virtual Reality (VR) devices. High-end Virtual Production (VP) projects use tracking technology to update the pose of a virtual camera twin, providing the studio team with immediate cinematography choices [1]. The twinned virtual camera scene render can replace green screen backgrounds electronically, or be displayed on physical LED backdrops and captured in-camera directly. Operators interact with the camera with almost no deviation from traditional handling.

We replicate the interaction between a movie camera and a user in a virtual environment, but consider a use case other than VP, specifically storyboarding. Our Camera Chameleon presents the virtual camera rendering on a small display attached to a motion-tracked device, such that the user manipulates something similar to a traditional camera, in real-time. This offers the potential to return the virtual cinematographer back to familiar tangible camera interaction, rather than relying on tools that are a poor fit for camera manipulation, such as a KM. While many task-specific peripherals have been studied to understand impact, there is little academic research into the effects of tangible, tracked interaction for virtual cinematography. Our research question is hence: **How do tangible virtual cameras affect a storyboarding task when compared with incumbent keyboard and mouse interaction?** We answer this through two contributions: 1) develop a user-friendly tangible cinematography device, and 2) perform the first rigorous user study into tangible device usage for virtual cinematography. The interaction and spatial cognition of tangible devices promises familiarity for traditional camera users. This has been an inspiration [2], [3] for the VP community. This motivates our work, and sets a deductive and empirical theoretical framework approach — to scrutinise how such devices perform against traditional peripherals.

## II. RELATED WORK

First, comparable approaches will be discussed; secondly, as relevant to the application, creativity measurements will be considered. Early tracked displays examples, such as the "*Chameleon*" concept [4], were primarily concerned with 3D navigation usability. Mouse interaction remained quicker than their 4-inch Chameleon device, but 3D perception was unaffected. Recent applications in virtual cinematography include "Screenplay Scout" [2] and "MobileVCS" [3] both developed hardware similar to our own, although lack of implementation detail is something we wish to address here and subject the concept to rigorous scrutiny. These ideas are now supported by products such as the Unreal game engine [5]. "Director's Lens" [6] extends this idea, combining a 6 DoF tracked display with a camera pose recommendation system. This is primarily driven by cinematography rules rather than any artistic model.

The combination of camera tracking, game engines and LED walls has led to VP [1], as demonstrated by shows such as The Mandalorian, and considered in pre-visualisation [7]. 3D visualisation interaction methods have been surveyed [8] and categorised into tactile, tangible, gesture, or hybrid.

Comparisons have been made between such approaches [9], with the Chameleon approach found to be slower than the

Work supported by EPSRC Centre for Digital Entertainment EP/L016540/1, EPSRC CAMERA 2.0 EP/T022523/1, and UKRI MyWorld Strength in Places SIPF00006/1

"*Pinch-Flick-Drag*" interaction used on smartphones. However, the difference reduced as a user gained experience and disappeared if the task included further interaction requirements. Tracked displays have also been studied in the context of map and document navigation [10]. In comparison to using a mouse, tracked devices have demonstrated comparable accuracy and increased speed [11], [12]. Users often provide positive feedback, in particular, that a tracked, tangible device accelerates the learning of interaction concepts. The success of tracked cameras as an interface inspires our development of a tangible camera-like device for virtual cinematography, which we name Camera Chameleon (CC). Development is described in Section III-C.

There are many approaches to measuring task performance. Being simple and quantitative, some studies consider accuracy and timing [11]. There is conflicting evidence for the speed differences between tracked and traditional peripherals [9], [11], [12] — it is task-dependent and influenced by prior user experience. As such, and from the gap in the literature on virtual cinematography applications, we wish to learn the difference also. We will measure speed, but not accuracy as it does not apply to storyboarding, a creative activity.

Creativity is critical throughout any production, particularly in earlier stages such as storyboarding; we wish to understand how a tracked camera affords the creativity of a user. To measure creativity, a definition is required [13] such as the *"the four P's of creativity"* [14] or *"creative tripod"* [15] frameworks. Various surveys [16], [17] have uncovered the disparity (or lack of) framework usage, indicating at the complexity. We address our aim with a combination of creative measurements, specifically protocol analysis, surveys and observations (see Section III-B).

## III. USER STUDY

To evidence the potential benefit of using our CC device we conducted a user study to compare it against KM peripherals. We designed the study with one task: moving a virtual camera to create a storyboard according to a provided script, within a time limit. This was selected to encourage exploration and creativity, plus reflect the real goals and pressures of film-making.

### A. Design

We chose a within-subjects study design, with a single independent variable of two conditions (KM and CC), presented in a counterbalanced order.

The KM condition involved sitting and using the KM peripherals to pilot a virtual camera in Unreal Engine (UE). A primary monitor displayed the first-person virtual camera.

The CC condition was done standing, with users moving themselves and the CC device with one or two hands. The first-person virtual camera feed was presented on the mini HDMI display of the CC, with movement being controlled by a combination of physical movement and gamepad joysticks.

For each condition, participants were asked to populate all 9 shots of a storyboard at least once, ideally 2–3 times, within 15 minutes. The script stayed the same throughout the experiment. Our storyboard interface was displayed on a secondary monitor in both conditions to allow the user to 'take' shots into the storyboard.

Upon completing a storyboard, the user could reset the camera position and storyboard to begin again. To restrain task complexity, the user had control of the camera pose only, with no lens, aperture, lighting, object or render control. No restriction was put on which of the four actors to focus on. The study required effort commensurate with a typical hours-worth of general office work.

### B. Measures

To understand user profiles, we implemented a pre-study questionnaire of seven 5-point Likert scale questions asking users to self-rate their experience in areas of film-making, photography and digital 3D environments. To understand task duration and movement trends, we record movement and timing data from the virtual camera, tracking handset, gamepad, and shortcut keypad at 10Hz. Captured shot images in all storyboards were stored at a resolution of $960 \times 540$ pixels. Thirty 7-point Intrinsic Motivation Inventory (IMI) [18] questions presented after each condition, recorded the users' experience in 5 subcategories of enjoyment, pressure, competence, usefulness and effort. Finally, twelve 5-point Likert scale questions asked for condition comparisons, and preferences for hypothetical future tasks. Unconstrained free-text fields were included, asking for feedback about KM, CC, usage and improvements. These measures enable quantitative objective analysis through numerical analysis of user timing, movement and survey responses. Thematic analysis of the free-text fields provides qualitative measures.

### C. Apparatus

Our hardware is a tangible user interface, according to the formal definition of Ishii et al. [19], and exhibits the four key characteristics [20]. We draw from previous research (see Section II) and development to utilise low-cost consumer VR tracking technology for our tangible CC. We imposed overall requirements of portability, low latency, and similarity to a real cinema camera. The CC rig has four key features:

1) **Absolute 6-DoF tracking**. Inspired by the success of graspable [21] and tangible interface [22] development, and considered necessary for realistic movement, absolute 6 DoF tracking is realised using a single Oculus Rift VR handset. The UE Blueprint graphical scripting system translates the absolute physical tracked movement of the handset onto the virtual twinned camera.
2) **Real-time display**. To support reactive exploration in a latency reduced system [23] , a Spout sender [1] and receiver [2] plugin renders the virtual camera feed onto a mini HDMI monitor, with low latency.
3) **6-DoF acceleration control**. A USB gamepad allows long-range movement, as otherwise a user is limited by

---

[1] OffworldLive Unreal Engine plugin https://offworld.live/.

[2] Spout video routing for Windows https://spout.zeal.co/.

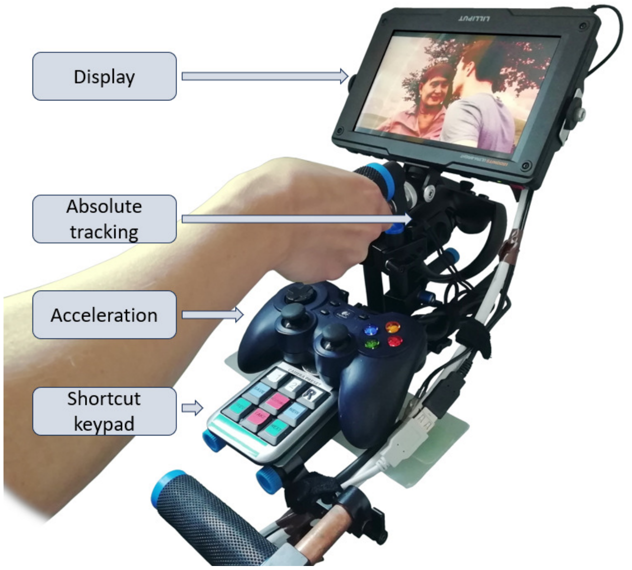

Fig. 1. Our tangible camera interface, referred to as *Camera Chameleon* (CC), with four main components: absolute tracking, display, acceleration, and keypad.

the tracked VR handset boundaries. Y-axis motion on the gamepad was translated into forward-facing camera motion in the virtual world.

4) **Keyboard shortcuts for rapid storyboard interaction**. To encourage completing the task within the 15-minute limit, time-saving shortcut keys were implemented. A wireless USB keypad was customized to enable rapid interaction with the storyboard user interface, as described further in Section III-D2.

A camera shoulder rig comprising 15mm aluminium tubing and brackets was modified to accept the above components. A top mounted handle enabled single handed holding, allowing a second hand to operate gamepad/keypad buttons. Figure 1 shows the final construction. The combined weight of the rig was 2.5kg, evenly distributed. HDMI, USB and 12v power cables were combined into a single 5m bundle, attaching the CC to a Windows PC.

### D. Visuals

*1) Virtual environment:* To provide adequate opportunity for exploration by participants a fictional scene was developed with a variety of settings that could reflect a range of film scenes. This includes a park with trees, a tower block building, street furniture and a skyline for the outdoor components. For the indoor setting, a living room is provided inside the building. Two groups of two standing (static) actors with neutral expressions were placed in the scene. Participants could move anywhere and focus on any of the (un-movable) actors.

*2) Storyboard UI:* We developed a Storyboard UI (see Figure 2), displayed on a large secondary monitor, and controlled by the CC wireless keypad. Keypad functions were assigned as: take shot, shot selection, camera preset and reset position. Before the study, we prepared a film script [24] of tense dialogue between two characters: the language was toned down and character names changed. It was separated into nine lines and each line pre-assigned to its own storyboard shot. Throughout the study, only one script was used, but participants could use any combination of four actors to focus on, with each new storyboard they created.

Fig. 2. A user in the study standing condition, manipulating CC, similar to a real camera. The storyboard UI is seen on the large screen in the background.

*3) Space Layout:* A desk was arranged at the front of a 3m × 4m space, with Oculus Rift sensors facing the user. Windows PC, primary monitor and KM were located on the desk. A large secondary (storyboard UI) monitor was positioned next to the desk. See Figure 2. A chair was provided for the KM condition, and moved away for the CC condition allowing freedom of movement.

### E. Participants

26 participants, optionally reporting age between 19 and 61 (*M*(24) = 33, *SD* = 11.7) and gender (8 female, 16 male) were recruited. Requirements were normal or corrected vision, and no previous experience of cinematography. No monetary compensation was offered and all participants provided consent.

### F. Procedure

Participants gave informed consent and completed a demographic questionnaire, followed by the two conditions. Each condition consisted of a 5-minute training session on the equipment, with the opportunity to practice with the system. Then, the participant was given 15 minutes to complete the storyboarding task according to the assigned condition. The user was verbally prompted when 5 minutes remained of the 15 minutes. Following that, the participant was asked to complete the 7-point scale IMI [18] questionnaire to obtain immediate feedback. The second, remaining condition was then tested using the same sequence as above; the entire procedure lasted approximately 1 hour. The study received ethical approval. The supplementary video shows the study in action.

## IV. EXPERT ASSESSMENT

An expert film-maker with 27 years of industry experience, having provided sound and mastering services to more than 200 feature films and commercials across Europe and Asia, and having directed commercials and drama films, was recruited to provide expert-level critique of the output storyboards. The expert was presented with a background to the project[3], and asked to remotely score each randomly presented storyboard via a web interface with three 101-point sliders, defaulting to 50.

Three sliders were chosen to balance reviewer workload, with the desire to produce quantitative review data. The 3 sliders related to components which are described as *Visually Pleasing*: Are the images pleasantly unexpected, convey beauty, or are different from 'the norm'?; *Readability*: Do the images convey the story appropriately? Would someone without seeing the script understand the 'jist' of the story?; and *Professionalism*: Are traditional / professional filming rules upheld (e.g. 180 degree rule, minimum angle violation etc.)?

## V. RESULTS

A total of 26 participants completed the study, each finishing both conditions. From the pre-study questionnaire results, most users reported low experience in most film-making and photography categories. Only casual photography and 3D experience contained a more balanced mix of people. From the latter we nominate two important groups to our study — 16 users with high amounts of digital 3D experience "hi3D", and 10 without "lo3D". "hi3D" is based on users' response as 'Agree' or 'Somewhat Agree' to the question of "I am experienced in navigating digital 3D environments (e.g. gaming, 3D animation, etc.)". "lo3D" users are nominated as those answering "Neither Disagree / Agree", "Somewhat Disagree" and "Disagree". Significance level is set at $\alpha = .05$.

### A. Quantitative

*1) Number of storyboards completed:* Depending on their individual speed, participants completed between 1 and 8 storyboards in each condition. Total completed storyboard count was 78 (KM) + 87 (CC) = 165, giving $165 \times 9 = 1485$ shots in total.

*2) Composition time:* A significant reduction ($t(25) = 2.61$, $p = .015$, Cohen's $d = 0.513$ ) with substantial effect size [25] in users' mean composition time (the duration of moving the camera and taking a shot) between KM ($M = 36.1$ seconds, $SD = 18.7$) and CC ($M = 27.8$ seconds, $SD = 9.5$) was observed. This indicates users of the CC complete more work compared to KM.

Analysing all 1485 composition times by their associated shot index shows duration is highest for shot index 0 — likely attributed to the initial navigation from the starting position. A slight increase in shot index 3 and 4 may be attributed to the particular script text — that of an actor screaming. There is also a common increase in composition time for the last shot — users may be reviewing all previous shots of the storyboard before finalising.

*3) Movement:* Novice (lo3d) users appear to move more using the CC than KM. We sum all absolute 6-DoF movements from all shots completed by each user, separated by condition. The mean fraction of total per-user camera movement favours the KM condition ($M = 55\%$, $SD = 11$) $t(25) = 23.81$, $p < .001$, Cohen's $d = 4.671$. A moderate positive correlation ($r(24) = 0.494$, $p = .010$) exists between a user's self-reported 3D experience, and the proportion of total movement being attributed to the KM condition.

When using the CC, users physically rotate more than using the gamepad for rotation. As a percentage of total rotation, tracked rotation dominates ($M = 77\%$, $SD = 11$) $t(25) = 35.61$, $p < .001$, Cohen's $d = 6.984$ compared to gamepad rotation. Conversely, as a percentage of total *translation* movement, the translation movement is dominated by gamepad usage ($M = 98\%$, $SD = 1$) $t(25) = 435.54$, $p < .001$, Cohen's $d = 85.416$ over tracked movement.

*4) IMI Results:* Here we outline the user feedback according to IMI questionnaires completed after each condition. We see a higher mean level of enjoyment with CC compared to KM of 0.69. Only enjoyment showed significant results between conditions $t(25) = -4.03$, $p < .001$, Cohen's $d = 0.79$. All other subcategories were not significant, with the lowest being effort ($p$=.332), as shown in Table I.

| | KM | | | CC | | | Paired | | |
|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *t*(25) | *M* | *SD* | *t*(25) | *M* | *p* | *d* |
| enjoyment↑ | 5.16 | 0.98 | 26.8 | **5.85** | 0.80 | 37.1 | -4.03 | <.001 | -0.790 |
| pressure↓ | 2.62 | 1.24 | 10.8 | 2.56 | 1.07 | 12.2 | 0.24 | .810 | 0.048 |
| competence↑ | 4.12 | 1.11 | 18.9 | 4.12 | 1.40 | 15.1 | -0.04 | .969 | -0.008 |
| usefulness↑ | 4.19 | 1.42 | 15.1 | 4.33 | 1.55 | 14.3 | -0.87 | .392 | -0.171 |
| effort↓ | 5.13 | 0.88 | 29.6 | 5.28 | 1.04 | 25.8 | -0.99 | .332 | -0.194 |

TABLE I

Statistics for the IMI results, obtained after each condition of KM and CC (min=1, max=7). n users=26. All $p < .001$ unless stated. $d$ is Cohen's $d$. **Bold** shows significant results. ↑: higher values are better. ↓: lower values are better

*5) Post-study questionnaires:* The questionnaires completed after the second condition/IMI, provided 5-point Likert scale responses to a range of metrics. The first rated the CC against KM on agree-to-disagree scales, such as being more preferable, natural, creative, and quicker. Strong agreement for creativity and naturalness was found, (see Table II ). All categories' means scored higher than 3, i.e. above "neither agree/disagree".

The second questionnaire asked the user what mix of peripherals would be preferable for future, hypothetical tasks. The five Likert options ranged from using only KM(1), through an equal mix of KM / CC (3), to only using CC(5). Only one question ("If I had to do this task all day") resulted in a mean score showing a slight preference towards KM ($M = 2.77$, $SD = 1.21$). All others showed preference towards CC, especially for creativity and training, with the highest being "If creativity

---

[3] Expert review online guide https://github.com/w00dw0rm/TVC_info

| Compared to KM, I find CC: | $M\uparrow$ | *SD* | *t*(25) |
|---|---|---|---|
| more preferable | 3.46 | 1.17 | 15.03 |
| more natural | 4.00 | 1.0 | 20.00 |
| more creative | 4.39 | 0.94 | 23.75 |
| quicker to achieve my goal | 3.54 | 1.07 | 16.91 |
| quicker to learn how to use | 3.20 | 1.18 | 14.00 |

TABLE II
RESULTS FROM A POST-STUDY 5-POINT LIKERT SCALE QUESTIONNAIRE, COMPARING CC TO KM, (1 = DISAGREE, 5 = AGREE). ALL $p < .001$. HIGHER RESULTS INDICATE A PREFERENCE TOWARDS CC.

was a goal" (*M* = 4.39, *SD* = 0.75) and "to train other people in Real cinematography" (*M* = 4.31, *SD* = 0.62)

### *B. Expert review of storyboard outputs*

We asked an expert film maker to rate the completed storyboards (annotated with the script text), randomly ordered, on 3 scales: visually pleasing, readability and professionalism. The expert took approximately 2 hours to complete the analysis of 168 storyboards. The ratings of 3 storyboards were removed due to errors, leaving 165 (78 KM and 87 CC) complete ratings. There is no significant difference in any of the three scores ($p > .083$), which we can say shows that the output quality is no different based on condition. We are encouraged by this parity: KM is a go-to method for professional use, and given the CC was novel to many participants, achieving similar quality is therefore good, with the added benefits of being quicker and more enjoyable.

Also, since the default values for the sliders were set to the mid-point (50), we acknowledge that it may have introduced a reviewer bias [26] to use that default value: the difference between the mean value and the default value for all of the three scores was less than 7.1. Interestingly, no significant correlation could be found between a user's 3D experience, and their output storyboard quality based on the three components tested: ($p > .323$).

### *C. Qualitative*

A thematic analysis was performed on the questionnaire free-text fields, to understand nuance of user perception not captured in quantitative analysis. A description-focused coding scheme was manually generated in an *inductive* fashion [27]. Codes were then grouped into themes. The themes identified positive and negative aspects. Benefits are drawn from the positive themes: 'supporting creativity', 'allowing refined movement', and 'feeling natural' — all of which contribute to encouraging answers to the research question of "How do tangible virtual cameras affect a storyboarding task when compared with incumbent keyboard and mouse interaction?" . Negative themes highlight areas for future improvements: 'weight' identified that participants noticed the difference compared to KM (a task not holding anything), but as a starting point, the mass of the CC device is considered a good starting point, being commensurate with professional camera equipment. 'Physical space' themes included comments about getting used to the act of moving physically, the desire to move more, and for two users the potential to become spatially confused when navigating the virtual environment. The themes drawn out here contribute to the overall impression of the CC device, and allow further investigations to progress the quality of tangible virtual camera peripherals.

## VI. DISCUSSION

Our results show users retain similar output quality when using KM, but at the same time are afforded several benefits. In support of creativity, levels of exploration and enjoyment are shown to be significantly increased. This is backed up by one user saying *"it was easier to stumble upon happy accidents using the tracked camera"*. Users also rated the CC as being significantly more creative and natural to use compared to KM. The time taken to compose each shot is shorter compared to KM use. A faster composition time for CC was not assumed: professional 3D animators are extremely quick using familiar KM devices, unlikely to be improved upon by physically moving a peripheral. However, for non-experts in our study (lo3D), the intuition of tangible devices clearly assists in the task completion rate.

We find that when using CC, most users physically rotate a lot, but only translate by small amounts. Large gamepad translation is required to move the camera from the starting position to the actors, so this may dominate the behaviour. Some users commented on the desire for a lighter-weight CC. This could be due to inexperience in holding cameras, and maybe compounded by the CC task being standing and the KM task being sitting. The 2.5kg mass of CC lies between consumer DSLR and professional cinema cameras: we believe the current device is a good starting point for such a peripheral.

Considering our original question of **How do tangible virtual cameras affect a storyboarding task when compared with incumbent keyboard and mouse interaction?**, we raise the following points. For educational purposes, the novelty, enjoyment, and speed at which users pick up the general concepts of moving a CC device can be utilised by students wishing to learn cinematography. Benefits may be realised in a collaborative classroom setting, without the isolation problems of VR headsets. For industry, virtual pre-production work using similar virtual environments and our CC hardware will be more efficient than traditional methods; saving money by using less time, or offering more iterations in the same time. The cost at which tracked cameras can be realised is also beneficial for budget-constrained projects that wish to make use of virtual technology. For the research community, we lay some foundations upon which further studies can be built , leading to further improvements in virtual camera interaction. Comparing CC to others is difficult because of the scarcity of previous results — something which we address in this work. **Future work**. We acknowledge the limited project scope, and encourage areas of further research on CC-like devices. These include lens controls and the influence of more "in-camera" controls, plus the ability to record dynamic trajectories rather than static images only. Future hardware development that addresses some users' concerns like weight and range of movement would also be of interest.

# REFERENCES

[1] M. Sawicki and J. Moody, *Filming the Fantastic with Virtual Technology: Filmmaking on the Digital Backlot*. Routledge, 2020.

[2] W. Bares, N. Buckner, and D. Castille, “Virtual camera scouting with 3dof motion trackers,” in *Proc. of the 48th Annual Southeast Regional Conf.*, 2010, pp. 1–2.

[3] G. Balakrishnan and P. Diefenbach, “Virtual cinematography: Beyond big studio production,” in *ACM SIGGRAPH 2013 Studio Talks*, New York, NY, USA, 2013.

[4] G. W. Fitzmaurice, S. Zhai, and M. H. Chignell, “Virtual reality for palmtop computers,” *ACM Transactions on Information Systems*, vol. 11, no. 3, pp. 197–218, 1993.

[5] U. Engine, “Live link vcam,” Sep 2021. [Online]. Available: https://apps.apple.com/us/app/live-link-vcam/id1547309663

[6] C. Lino, M. Christie, R. Ranon, and W. Bares, “The director’s lens: an intelligent assistant for virtual cinematography,” in *Proc. of the 19th ACM international conf. on Multimedia*, 2011, pp. 323–332.

[7] T. Muender, G. Volkmar, D. Wenig, and R. Malaka, “Analysis of previsualization tasks for animation, film and theater,” in *Ext. Abstracts of the 2019 CHI Conf. on Human Factors in Computing Systems*, 2019, pp. 1–6.

[8] L. Besançon, A. Ynnerman, D. F. Keefe, L. Yu, and T. Isenberg, “The state of the art of spatial interfaces for 3d visualization,” in *Computer Graphics Forum*, vol. 40, no. 1. Wiley Online Library, 2021, pp. 293–326.

[9] M. Pahud, K. Hinckley, S. Iqbal, A. Sellen, and B. Buxton, “Toward compound navigation tasks on mobiles via spatial manipulation,” in *Proc. of the 15th int. conf. on Human-computer interaction with mobile devices and services*, 2013, pp. 113–122.

[10] K. Yee, “Peephole displays: pen interaction on spatially aware handheld computers,” in *Proc. of the SIGCHI conf. on human factors in computing systems*, 2003, pp. 1–8.

[11] L. Besançon, P. Issartel, M. Ammi, and T. Isenberg, “Mouse, tactile, and tangible input for 3d manipulation,” in *Proc. of the 2017 CHI conf. on human factors in computing systems*, 2017, pp. 4727–4740.

[12] P. Song, W. Goh, C. Fu, Q. Meng, and P. Heng, “Wysiwyf: exploring and annotating volume data with a tangible handheld device,” in *Proc. of the SIGCHI conf. on human factors in comp. sys.*, 2011, pp. 1333–1342.

[13] M. Batey, “The measurement of creativity: From definitional consensus to the introduction of a new heuristic framework,” *Creativity Research Journal*, vol. 24, no. 1, pp. 55–65, 2012.

[14] M. Rhodes, “An analysis of creativity,” *The Phi delta kappan*, vol. 42, no. 7, pp. 305–310, 1961.

[15] S. Colton, “Creativity versus the perception of creativity in computational systems.” in *AAAI spring symposium: creative intelligent systems*, vol. 8, 2008, p. 7.

[16] P. Karimi, K. Grace, M. Maher, and N. Davis, “Evaluating creativity in computational co-creative systems,” *arXiv preprint arXiv:1807.09886*, 2018.

[17] A. Jordanous, “Evaluating evaluation: Assessing progress in computational creativity research,” 2011.

[18] E. McAuley, T. Duncan, and V. V. Tammen, “Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis,” *Research quarterly for exercise and sport*, vol. 60, no. 1, pp. 48–58, 1989.

[19] H. Ishii, “Tangible bits: beyond pixels,” in *Proc. of the 2nd international conf. on Tangible and embedded interaction*, 2008, pp. xv–xxv.

[20] B. Ullmer and H. Ishii, “Emerging frameworks for tangible user interfaces,” *IBM systems journal*, vol. 39, no. 3.4, pp. 915–931, 2000.

[21] G. W. Fitzmaurice, H. Ishii, and W. A. Buxton, “Bricks: laying the foundations for graspable user interfaces,” in *Proc. of the SIGCHI conf. on Human factors in computing systems*, 1995, pp. 442–449.

[22] D. Harley, A. P. Tarun, D. Germinario, and A. Mazalek, “Tangible vr: Diegetic tangible objects for virtual reality narratives,” in *Proc. of the 2017 conf. on designing interactive systems*, 2017, pp. 1253–1263.

[23] P. Caserman, M. Martinussen, and S. Göbel, “Effects of end-to-end latency on user experience and performance in immersive virtual reality applications,” in *Entertainment Computing and Serious Games: First IFIP TC 14 Joint Int. Conf., ICEC-JCSG*. Springer, 2019, pp. 57–69.

[24] Q. Tarantino, *Reservoir Dogs (script)*. Miramax, 1992.

[25] J. Cohen, *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

[26] M. Liu and F. G. Conrad, “Where should I start? on default values for slider questions in web surveys,” *Social Science Comp. Rev.*, vol. 37, no. 2, pp. 248–269, 2019.

[27] V. Braun and V. Clarke, “Using thematic analysis in psychology,” *Qualitative research in psychology*, vol. 3, no. 2, pp. 77–101, 2006.