Interdisciplinary Practice in Industrial Design, Vol. 222, 2026, 144–155
https://doi.org/10.54941/ahfe1007754

Photo Presentation Methods for Designing a Consistent Photography Experience

Yusei Sakata and Wonseok Yang

Graduate School of Shibaura Institute of Technology, Koto-ku, Tokyo 135-8548, Japan

ABSTRACT

With the generalization of Multi-device usage, photographic images are captured by diverse devices, yet their viewing is consolidated into integrated smartphone applications. However, in the pursuit of efficiency and data integration, the unique characteristics of capture devices—such as physical tactile feedback, operational processes, and the atmospheric context physically experienced by users—are often discarded, treating image data merely as homogeneous pixel information. Consequently, the context of the capturing moment and device-specific uniqueness are diluted during the viewing stage, compromising the consistency between the capturing and viewing experiences. This research examined the UI elements designed to reconstruct the unique experiential value of capture devices within viewing software by organically connecting both experiences. Through a survey of UI elements in existing image management services and text mining of user reviews regarding various capture devices, we revealed that users value not only functional efficiency but also physical constraints and the physical operation itself. Based on these findings, we investigated the characteristics required for the viewing experience using a framework of "Materiality, Context, and Embodiment," focusing on reconstructing physical presence in digital environments. The results indicate that imparting sensory characteristics, such as physical presence and tactile feedback associated with operations, is crucial for evoking memories and sensations of the capturing moment, even in digital spaces. These elements are identified as effective means to ensure consistency between capturing and viewing experiences, complementing the context and uniqueness, otherwise diluted in digital environments.

Keywords: Seamless UX, Multi-device, Interface design

INTRODUCTION

This research focuses on the issue where consistency is compromised between the capturing experience using physically tangible "capture devices" and the viewing experience on integrated digital applications, examining UI elements to organically connect the two. Photographic images are captured using diverse hardware, such as smartphones, digital cameras, and wearable devices. In this research, hardware possessing the function of mediating the act of capture is collectively defined as "capture devices." Each capture device possesses a unique "hardware experience such as the tactile feedback of releasing the shutter, physical gestures of looking through a viewfinder, or waiting

Received February 19, 2026; Revised April 11, 2026; Accepted April 26, 2026; Available online July 20, 2026

time for development. However, in the current photography environment, captured data are consolidated into cloud services or OS standard integrated applications, and all image data are treated as homogeneous pixel information, regardless of the source device. In this process, the device-specific context physically felt by the user during capture is discarded, making it difficult to evoke it during the viewing stage. The "dilution of experience because of data aggregation" is a critical issue in digital photography research. Rodden and Wood pointed out that in environments where massive amounts of data are accumulated, organization and rediscovery become difficult, hindering their review and utilization (Rodden & Wood, 2003). Odom et al. also demonstrated that designs promoting revisitation and recollection through approaches that differ in efficiency can induce behavioral changes in users (Odom et al., 2014). However, while these existing studies focus on promoting review or realizing intuitive operability, the perspective of designing coordination between hardware and software experiences—specifically, the elements of physical and sensory experiences generated by specific capture devices that can be amplified and evoked within the viewing experience—has not been sufficiently organized. Therefore, this research analyzes the challenges faced by current integrated image applications that pursue efficiency through a survey of UI elements in existing services and text mining of user reviews, and examines UI elements to connect the hardware experience during capture and the software experience during viewing as a consistent entity.

Lack of "Efficiency" and "Consistency of Experience" in Digital Photography

The development of digital photography has dramatically streamlined the process from capture to sharing. Kirk et al. defined the lifecycle of digital photography using the "PhotoWork" model, focusing on the efficiency of post-capture tasks such as "downloading, organizing, editing, and sharing" (Kirk et al., 2006) (see Figure 1). This research focuses on the disconnect that has emerged between the "Capture" phase and the "File picture" (organizing and viewing) phase within this lifecycle model, which should inherently be closely related. Current integrated applications on smartphones pursue this efficiency to an extreme extent by arranging vast numbers of photographic images as homogeneous thumbnails in a grid to enhance the integration of searching and browsing. However, in this process of efficiency, photographic images tend to be separated from the "context" of the capturing moment and the device-specific "operational process," treated merely as pixel information. Broekhuijsen et al. argue that a shift in perspective is necessary from such "PhotoWork" to "PhotoUse," which emphasizes the long-term meaning and context of photographic images (Broekhuijsen et al., 2017).

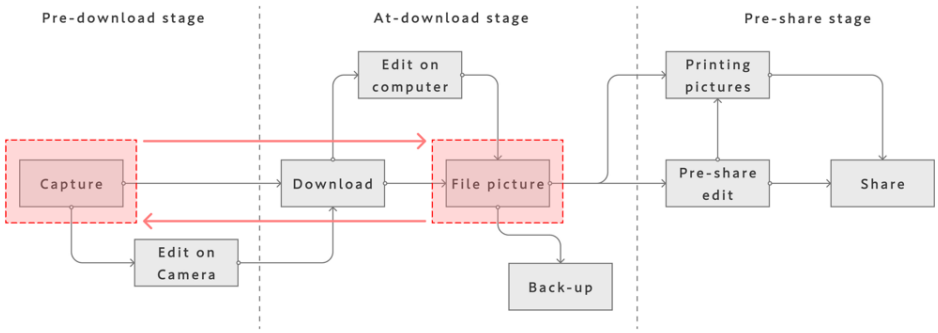

Figure 1: The lifecycle of digital photography (Kirk et al., 2006) and the scope of this research.

However, analog experiential values, which stand in contrast to such efficiency, are being reevaluated. Giaccardi and Karana proposed the concept of "Digital Materiality," suggesting that data can evoke a sense of touch similar to physical materials through specific forms of expression and manipulation (Giaccardi & Karana, 2015). Caoduro and Gulotta note that expressions of nostalgia and aging in digital environments can be significant elements in forming users' emotional attachments (Caoduro, 2014; Gulotta et al., 2013). In fact, according to research by Vukojević et al., younger generations, particularly Gen Z, find value in retro technologies like Polaroid not merely out of nostalgia but for their "tangibility" and "imperfection" (Vukojević & Castagnola Zamudio, 2025). Particularly important is the finding of Petrelli et al. in their research on "Digital Mementos," which showed that users feel a "material presence" absent in digital data through the touch of physical photographs (prints) and photographic equipment (Petrelli et al., 2013). The challenge is that the homogeneous viewing experience in current integrated apps has discarded these device-specific "experiential values."

Consistency Between Hardware Experience (Capturing) and Software Experience (Viewing)

To ensure consistency between the capturing and viewing experiences, it is necessary to understand the behavioral process of how users perceive devices and initiate operations. In this research, the "Framework for Designing Intuitive Interfaces" by Hirokawa et al. is adopted as the fundamental theory (Hirokawa & Ohno, 2014) (see Figure 2). Hirokawa et al. argue that intuitive usage becomes possible when the three elements of "Form," "Meaning," and "Operation" in the flow of user action execution are integrated through "Familiarity" based on user memory and "Organization" that structures these elements.

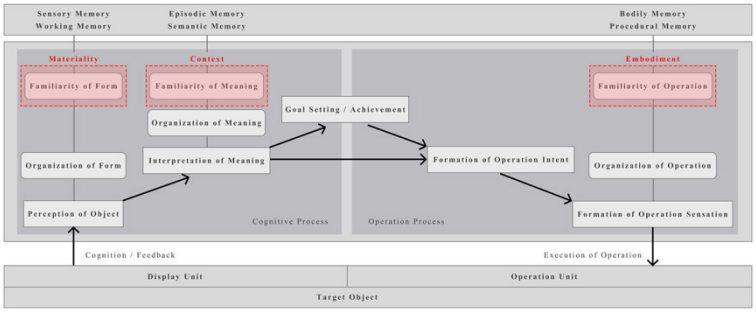

Figure 2: Organization and familiarity in the flow of action execution (Hirokawa & Ohno, 2014) and the approach of this research.

According to this framework, many existing interfaces utilize metaphors to ensure Familiarity of Form and facilitate the Organization of Meaning, while achieving mastery through the Organization of Operation. However, Hirokawa et al.'s model ideally aims to improve transparency and usability, making the interface unnoticeable. In viewing photographic images, simply being easy to use with high transparency is insufficient to evoke the rich experience of the capturing moment. The value of the capturing experience exists within certain "physical burdens" and "constraints," such as the tactile feedback when releasing the shutter or the waiting time for development. While the grid display in current integrated apps excels in the Organization of Form regarding visibility, it has discarded the context and embodiment of the capturing moment. Therefore, while the approach of this research is based on the "Form, Meaning, and Operation" framework proposed by Hirokawa et al., we extend it by incorporating new perspectives of "Materiality," "Context," and "Embodiment" to reconstruct the quality of the capturing experience within the software. By utilizing this extended framework, we examine specific UI elements to reconstruct not merely the reproduction of operations, but the "quality of experience"—such as the capture context and device-specific tactile feedback—on the viewing software.

RESEARCH METHOD

In this research, to derive UI elements that ensure consistency between the capturing and viewing experiences, we employed a two-step approach. First, we conducted a survey of UI elements in existing services and brainstorming to redefine a new framework for evaluating experiences that transcends mere functionality. Next, using this framework, we performed text mining of user reviews to extract the experiential value unique to each device and examined the elements that connect both experiences (see Figure 3).

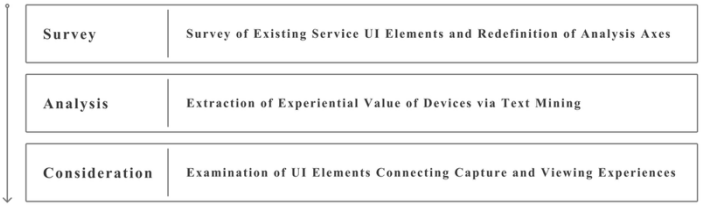

Figure 3: Research procedure.

1) Survey of Existing Service UI Elements and Redefinition of Analysis Axes

- Survey of UI Elements and Identification of Issues

To clarify how device characteristics and capturing experience are preserved, and how consistency between capturing and viewing is maintained in current image management applications, we conducted a survey of UI elements targeting 20 major services. We selected standard OS apps, camera manufacturer apps, and third-party apps that attempt to reproduce the experience as survey targets (see Figure 4). For classification, to comprehensively grasp the interface not merely as a collection of functions but as a series of experiences ranging from visual metaphors to physical operations, we used the three elements of intuitive interfaces proposed by Hirokawa et al. (2014) (see Figure 5). We classified and organized how these elements are implemented in each app, and extracted issues from the perspective of ensuring consistency in a modern Multi-device environment.

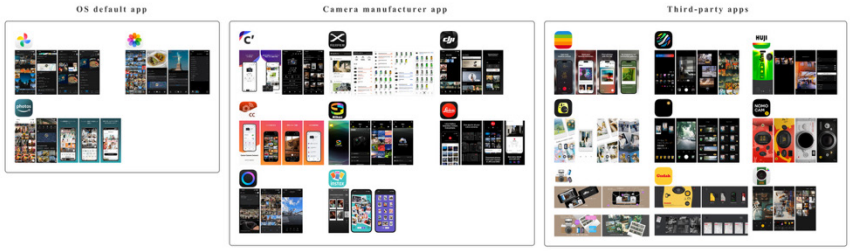

Figure 4: Surveyed services.

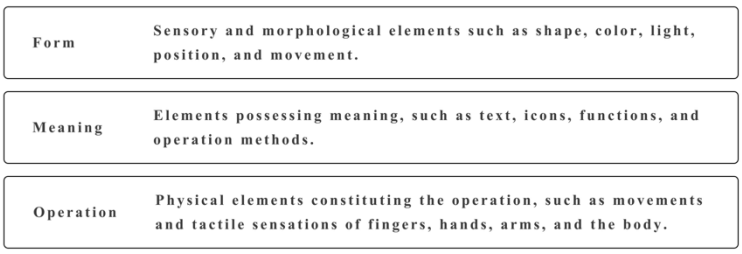

Figure 5: Three elements of "form, meaning, and operation" by Hirokawa et al. (2014).

- Extraction of Elements via Brainstorming and Redefinition of Analysis Axes

To structure "experiential values" that cannot be captured by the functional classification of the UI elements extracted in the above survey, we conducted a qualitative analysis using brainstorming and the KJ method with six university students studying designs. This process aims to verify the impact of individual UI elements on user emotions and bodies, and systematize them as new analysis axes. The specific procedure is as follows.

Step 1: Extraction of Elements: From the surveyed applications, we comprehensively extracted and carded the presentation elements implemented independently of functional necessity and characteristic UI expressions that were difficult to classify using Hirokawa et al.'s framework.
Step 2: Qualitative Grouping: We conducted brainstorming on the extracted elements, focusing not on "function" but on the "sensations" and "psychological effects" received by the user. We grouped (clustered) elements with similar experiential values and organized their structure.
Step 3: Definition of Concepts: We abstracted the characteristics of each formed group and defined them as new analysis axes to reproduce the hardware experience on software.

2) Extraction of Experiential Value of Devices via Text Mining

- Data Collection

To derive specific design requirements based on the new analysis axes, we quantitatively analyzed how users perceive and evaluate each capture device. As data sources, we primarily used customer reviews from "Amazon," where honest purchase evaluations are gathered, and also included articles from the media platform "note" to supplement more detailed experience episodes and emotional contexts. We set five categories with different qualities of experience as analysis targets and collected 200 review data points (40 for each device) (see Figure 6).

Figure 6: Target devices.

- Analysis Method

We performed morphological analysis on the collected text data and then conducted co-occurrence network analysis using the quantitative text analysis tool "KH Coder." In the analysis, we focused on the co-occurrence

relationship between "adjectives" representing the quality of experience and "verbs" representing physical actions among the extracted vocabulary. Through this network structure, we visualized the type of "experiential value structure" users have for each device and attempted to extract elements to connect the hardware and software experiences.

RESEARCH RESULT AND CONSIDERATION

1) Results of Existing Service UI Element Survey and Redefinition of Analysis Axes

- Trends and Classification of UI Elements in Existing Services

We classified and organized the UI elements of 20 major image management services using the framework of "Form, Meaning, and Operation" by Hirokawa et al. (2014). The results confirmed clear differences in UI design approaches depending on the service provider and purpose (see Table 1).

Table 1: UI element results.

	UI Elements
Form	Grid display, 3D models of capture equipment, Film/Photographic paper frames, Body/Lens graphics, Brand logos/Colors, Viewfinder UI, Analog counter
Meaning	Name, Metadata, Date and time, Location information, Subject, Media type, Event/Album name, Film simulation history, Development timer (Waiting time), Icons (Favorites)
Operation	Transfer/Import operations, Film winding, Shaking the device, Mimicking physical dials, Pinch-in/out (NUI), Swipe (Print ejection), Physical button integration

In applications developed by camera manufacturers and certain third parties, intentional designs aimed at preserving the context of the capturing moment were observed. These applications utilize elements of "Form," "Meaning," and "Operation" to visually reproduce hardware and incorporate physical actions, attempting to connect capturing gestures to the viewing experience. In contrast, OS standard integrated applications tend to exclude these elements to the greatest extent possible. By arranging photographic images as homogeneous thumbnails in a grid, the physical characteristics and operational embodiment of each capture device are discarded. As a result, it became clear that in these integrated apps, the context of the capturing moment is easily disconnected, and images are treated merely as homogeneous pixel information separated from the physical experience.

- Extraction of Limitations and Issues of the Existing Framework

Analyzing existing services using Hirokawa et al.'s framework (2014) revealed limitations in evaluating the "reconstruction of capturing experience." First is the "ambiguity of boundaries between elements." In modern GUIs, visual elements often embody meaning and serve simultaneously as operation triggers, making it difficult to analyze them separately. Second is the "lack of experiential value." While the framework effectively evaluates efficiency and usability, it is insufficient for assessing the recollection of the capturing experience. Elements such as "waiting time" or "noise" are often excluded as

inefficient from a functional perspective, making it impossible to appropriately classify design intents aimed at fostering attachment.

- Extraction of Elements via Brainstorming and Redefinition of Analysis Axes

Based on the issues identified above, we focused on elements treated as "inefficiency" or "noise" in the existing framework and conducted a qualitative analysis using brainstorming. As a result, the following three new axes were extracted as components not for functional "convenience" but for reproducing the "experience" of the capturing moment (see Figure 7).

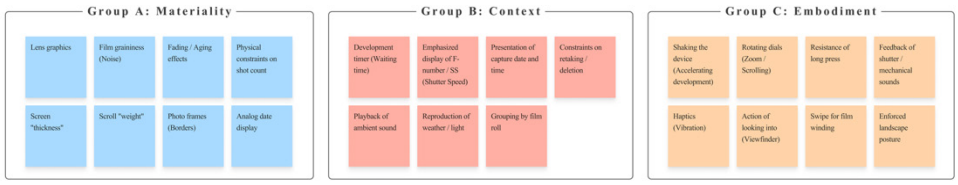

Figure 7: Grouping of UI elements via brainstorming and derivation of three elements.

① **Materiality** Defined as elements that give physical presence and media characteristics to digital data, such as device-specific textures, noise, light leaks, or aging effects, going beyond visual layout information.
② **Context** Defined as elements that extend beyond the display of mere metadata to include the invisible narrative surrounding a single image, such as the intent at the time of capture, the environmental sound of the scene, and the flow of time before and after.
③ **Embodiment** Defined as ritualistic elements that evoke the physical sensations of the capturing moment through operations accompanied by physical burdens—such as weight, resistance, and posture—rather than unconscious, smooth operations.

2) Results of Extraction of Experiential Value of Devices via Text Mining
The analysis of vocabulary connections and frequent word trends revealed distinct experiential value structures for each device group (see Figure 8).

- Co-occurrence Network Analysis

First, in instant and film cameras, functional "inconvenience" and physical "effort" are accepted as positive values. For instant cameras, the photograph as a physical "object" promotes on-the-spot "sharing" and "self-expression," while for film cameras, constraints like the inability to retake shots and the "delay" until development elevate the obtained results into "special" memories. Next, mirrorless cameras and smart glasses both generate high immersion as extensions of the body. In mirrorless cameras, physical operations of the "viewfinder" and "dials," along with shutter "sound," enhance active "immersion." Conversely, in smart glasses, "hands-free" and "voice control" features make the device transparent, evaluated for their ability to record "moments" without interrupting the flow of "life." On the other hand, the greatest values of smartphones are their overwhelming "physical proximity" and "efficiency." Being always in a "pocket," they record "casual"

daily moments in "large quantities" for immediate "sharing"; however, the vocabulary related to "attachment to inconvenience" or "physical tactile feedback" found in other devices is scarce.

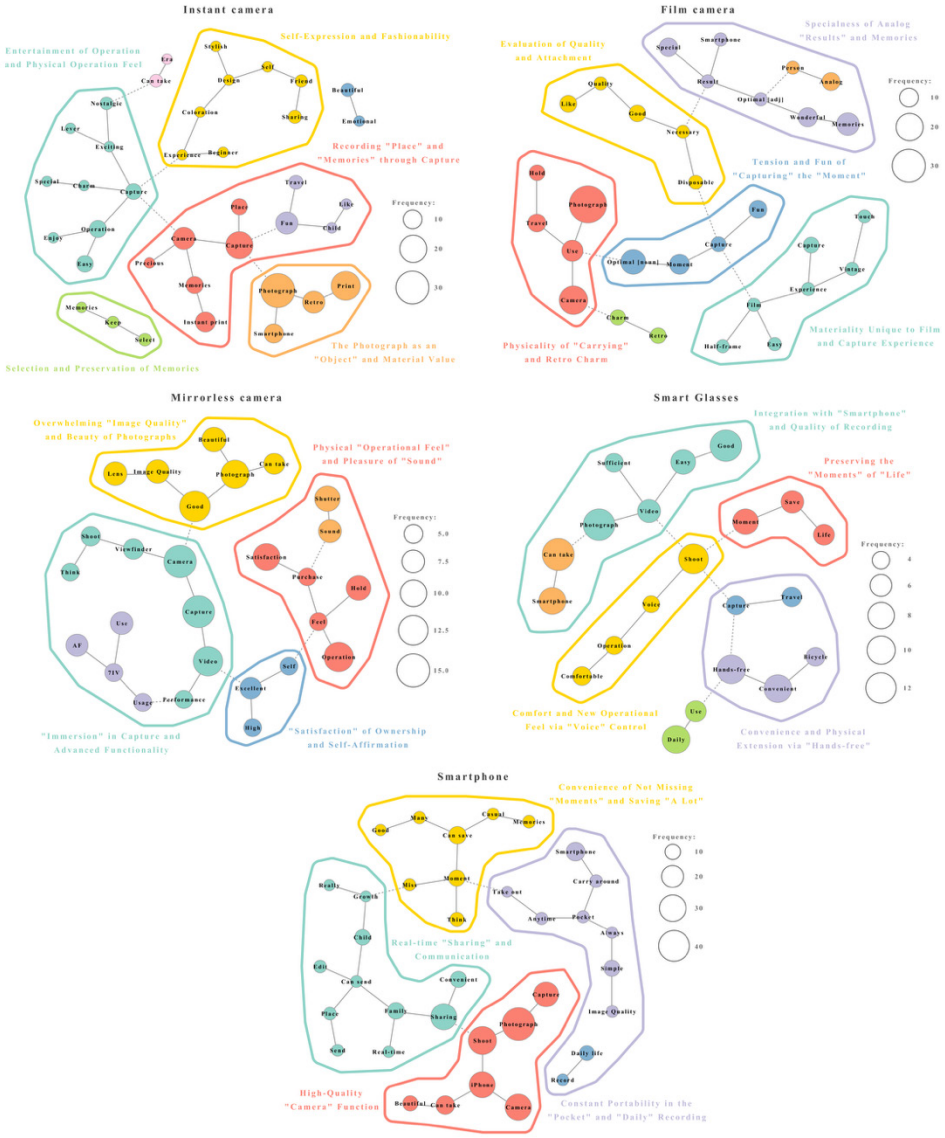

Figure 8: Co-occurrence network.

- Correspondence Analysis of Devices and Extracted Words

Furthermore, a correspondence analysis was conducted to overview the relationship between devices and extracted words (see Figure 9). The results structured experiential values along two axes: "Daily Efficiency vs. Intentional Capture Act" and "Material/Emotional Value vs. Functional/Physical Extension Value." In this distribution, smartphones are positioned at the pole of "efficiency," showing a statistically clear distance from the

"material/emotional" domain of instant and film cameras and the "active operational feel" of mirrorless cameras. This objectively demonstrates that "Materiality" and "Embodiment" are structurally missing in the current smartphone-centric viewing experience, underscoring the necessity of integrating these elements for its reconstruction.

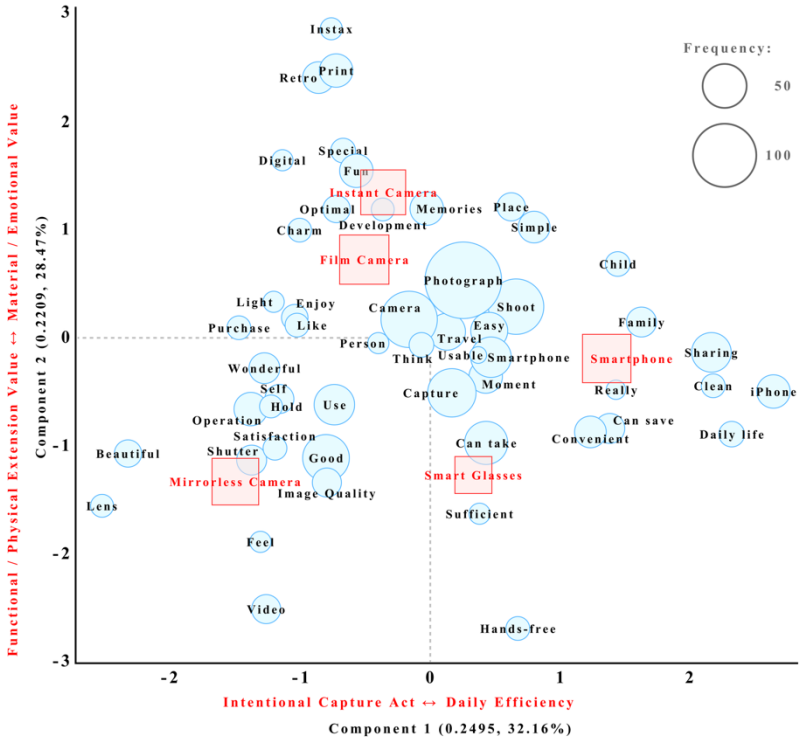

Figure 9: Results of correspondence analysis of devices and extracted ords.

CONCLUSION

In this research, addressing the issue where the consistency of the experience from capturing to viewing in a Multi-device environment is compromised, we examined elements to reconstruct the experiential value unique to capture devices at the viewing stage. The survey of UI elements in existing image management services and text mining of user reviews confirmed that even in a digital environment, users find strong value in analog "Materiality," the "Context" at the time of capture, and the "Embodiment" associated with operation. Based on these findings, the analysis indicates in designing the viewing experience, it is effective to impart physical and sensory characteristics—such as the presence of the object and tactile feedback of operation, in addition to efficient access. Specifically, characteristics such as the materialization of information by introducing intentional delays or weights to timelines otherwise consumed infinitely, the reconstruction of context by endowing single image data (often treated merely as pixel information) with narratives, such as ambient sound and the flow of time, and ritualistic

operations requiring active physical intervention beyond mere visual confirmation were extracted as primary factors contributing to user memory recollection and the formation of attachment to tools. These characteristics serve as vital clues for restoring the "quality of experience" in the digital space, which has heretofore been discarded as a priority for efficiency and data integration. The perspective based on "Materiality, Context, and Embodiment" presented in this research organically connects the capturing and viewing experiences, providing a new guideline for reconstructing the value of memory in digital photography and the relationship with capture devices.

REFERENCES

Banks, R., & Sellen, A. (2009). Shoebox: Mixing storage and display of digital images in the home. Proceedings of the 3rd International Conference on Tangible and Embedded Interaction (TEI '09), 135–142.

Broekhuijsen, M., van den Hoven, E., & Markopoulos, P. (2017). From PhotoWork to PhotoUse: Exploring personal digital photo activities. Behaviour & Information Technology, 36(7), 742–759.

Caoduro, E. (2014). Photo filtering apps: Digital nostalgia and the materiality of the photographic image. Networking Knowledge: Journal of the MeCCSA Postgraduate Network, 7(2).

Dourish, P. (2001). Where the Action Is: The Foundations of Embodied Interaction. MIT Press.

Giaccardi, E., & Karana, E. (2015). Foundations of materials experience: An approach for HCI. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15), 2447–2456.

Golsteijn, C., & van den Hoven, E. (2013). Facilitating parent-teenager communication through interactive photo cubes. Personal and Ubiquitous Computing, 17(2), 233–246.

Gulotta, R., Forlizzi, J., & Faste, H. (2013). Digital systems and the material of legacy. Proceedings of the 2013 Conference on Designing Interactive Systems (DIS '13), 633–642.

Hirokawa, M., & Ohno, T. (2014). A Framework for Designing Intuitive Interfaces based on the Organization and Familiarity of Actions. Transactions of the Virtual Reality Society of Japan, 19(2), 247–259.

Kamiyama, T., & Nakamura, S. (2015). Investigation on availability of photographing behavior for searching subjective photo lifelog. IPSJ SIG Technical Report, 2015-HCI-165(13), 1–8.

Kirk, D. S., Sellen, A. J., Rother, C., & Wood, K. R. (2006). Understanding photowork. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06), 761–770.

Kirk, D. S., Sellen, A., & Harper, R. (2010). Opening up the family archive. Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10), 261–270.

Odom, W., Banks, R., Durrant, A., Kirk, D., & Pierce, J. (2012). Photobox: On the design of a slow technology. Proceedings of the Designing Interactive Systems Conference (DIS '12), 665–668.

Odom, W., Sellen, A., Harper, R., & Thereska, E. (2014). Designing for slowness, anticipation and re-visitation: A long term field research of the photobox. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14), 1961–1970.

O'Hara, K., Harper, R., Mentis, H., Sellen, A., & Taylor, A. (2012). Food for Talk: Phototalk in the context of sharing a meal. Human–Computer Interaction, 27(1-2), 124–150.

Petrelli, D., Hoven, E. van den, & Whittaker, S. (2013). Making history: Intentional capture of future memories. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13), 1723–1732.

Rodden, K., & Wood, K. R. (2003). How do people manage their digital photographs? Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03), 409–416.

Stevens, M., Abowd, G. D., Truong, K. N., & Vollmer, F. (2003). Getting into the Living Memory Box: Family archives & holistic design. Personal and Ubiquitous Computing, 7(3-4), 210–216.

Vukojević, J., & Castagnola Zamudio, P. G. M. (2025). Beyond Nostalgia: Understanding Gen Z's Motivations for Using Polaroid Cameras (Master's thesis). Jönköping University.