A Cross-Media Modal Study of Painting Images in the Context of New Media- From Visual Observation to Multi-Sensory Immersion

Wei Guo

Film, Animation Arts Management,

The Graduate School, Sejong University. South Korea.

E-mail: 15234189097@163.com.

Abstract

In the current era, where digital media deeply intervene in artistic creation, the spread and experience of painting images are being redefined. Research on painting images has often focused on form and style, while neglecting the role of the medium in perception and meaning generation. This study, based on the perspective of media modality, re-examines painting images, transcending traditional frameworks of art history or iconography. This research not only places painting images within the overall framework of media modalities, systematically analyzing their material, sensory, and spatial properties, but also introduces visual perception theory to explore the perceptual mechanisms and psychological guidance effects of painting images across different media forms. Through the dual perspectives of media modality and visual perception, this study focuses on the visual language of Vincent van Gogh's paintings, selecting the film *Loving Vincent*, the virtual reality work *The Night Café VR*, and the immersive exhibition: *An Gogh: The Immersive Experience* as research subjects. It explores the mediation mechanism of painting images in cross-media modalities, with a particular focus on how the migration of media facilitates the re-generation of meaning in painting images and the reorganization of audience cognition, thereby revealing the diversified artistic language and new aesthetic interaction modes of painting images in the context of new media.

Keywords: Painting images; Lars Estlund; Media modality; Visual perception theory; Arnheim; Vincent van Gogh

1.Introduction

With the rapid development of new media technologies, the mode of image information spread has shifted from traditional one-way output to a two-way interactive communication process. This change has not only broadened the dimensions in presenting but has also significantly enhanced the active role of the audience in the art experience. Audiences have transitioned from passive appreciators to active participants, which has consequently redefined their aesthetic approaches and perceptual pathways. In this media context, painting art, in order

to adapt to the evolution of the times and the innovation of communication methods, has started to actively incorporate new media techniques, reshaping visual experiences through the extension of real space and the construction of virtual space. The visuality and expressive forms of traditional painting have gradually evolved from the previous "visual observation" to "multi-sensory immersion." Artistic creation no longer solely relies on flat visual language, but instead uses images, sound, space, and interactive technologies to form a complex media perceptual structure.

Various art projects that utilize new media as a medium have emerged, focusing on the contemporary expression of Vincent van Gogh's painting visual language. These works not only reshape van Gogh's artistic image but also reflect the process of migration and reconstruction of painting images across different media. The film *Loving Vincent*, the virtual reality work *The Night Café VR*, and the immersive exhibition *Van Gogh:The Immersive Experience* are representative practices of this media shift. These three art forms, through techniques such as animated images, virtual reality, and immersive multimedia spaces, extend van Gogh's visual language from a two-dimensional canvas to dynamic video spaces and interactive virtual domains, constituting a reinterpretation and regeneration of the traditional semantics of painting. They demonstrate how painting, in the context of new media, achieves the migration and reconstruction of its visual language through media transformation—continuing the emotional intensity and color expression of van Gogh's works, while reshaping time and space and updating perceptual modes, allowing the audience to enter the artist's spiritual world in a completely new immersive way.

In the prior research on artistic projects related to the life and works of Vincent van Gogh, studies on the animated film *Loving Vincent* can be roughly divided into three areas. First, from the perspective of film literature, research explores the narrative features of the film (Tian&Zhang,2019). Second, from the angle of animation techniques, most studies focus on how the film breaks traditional narrative structures and visual conventions through experimental animation, thus subverting the audience's habitual cognitive patterns (Lee,2024). Third, from the perspective of animation's artistic image characteristics, research examines how van Gogh's pastoral works from his Arles period are integrated with animated scenes to achieve dynamic representations that transcend time, space, and reality (Yao,2018).

In the research on *The Night Café VR,*scholars have continuously deepened their studies on the narrative mechanisms, spatial construction, and user experience in VR animation and film (Guo, 2025; Wang, 2018; Yan, 2024). Discussions on van Gogh's immersive exhibitions mainly focus on the following aspects: the aesthetic innovation from the fusion of art and technology, where digital media expands the expressive space of traditional painting, shifting the art experience from visual appreciation to full sensory participation (Sainae&Jang,2021);

the relationship between art consumption and the experience economy, with immersive exhibitions becoming an important form of contemporary cultural industries, promoting art popularization while also triggering the critique of "de-artification" of art (Mondloch & Kate, 2022); and the reconstruction of the relationship between the digital halo and the audience, where the art viewing experience in the new media environment reshapes the presence of art, transforming the audience from passive appreciators to co-creators of the experience (Baranseli & Ebru Selcan, 2024).

Based on the findings from the previous research, in cross-media expression there is also a relative lack of studies on the overall transformation of artistic projects related to van Gogh's paintings, especially in the context of the change in media modality. Specifically, how the characteristics of painting images are mediated through media modalities during media migration, thereby achieving the re-generation of meaning in painting images and the reorganization of the audience's cognitive structure.

Therefore, this study proposes the following three research questions:

1.How does the change in media modality affect the formal expression of painting images?

2.In the context of new media, what are the characteristics in those elements influencing the conversion of media modality in the migration of painting images?

3.How do different media in van Gogh's works affect the cognitive construction of the audience?

2.Media Modality Theory and Visual Perception

Based on Rudolf Arnheim's theory of visual perception, this study conceptualizes and analyzes the ten elements he categorized, examining the conventional animation framework expressed by each element. Audiences process information about light, color, space, and motion characteristics in painting images through their visual systems. These visual features not only construct the sense of depth and spatiality of the image but also trigger emotional responses and focus the audience's attention, thereby enhancing the depth and resonance of the immersive experience. Meanwhile, Lars Elleström's cross-media model, based on the concept of multimodality, analyzes the media mediation mechanism in the process of reinterpreting van Gogh's painting images. Elleström's model distinguishes four media modalities: material modality, spatiotemporal modality, sensorial modality, and semiotic modality. All media products are realized through some materiality, such as flat, raised, or solid forms, and experienced in spatiotemporal modes like two-dimensional, three-dimensional, continuous, or discontinuous, as well as perceptual modes like auditory or visual. The semiotic modality originates from the perception of material, spatiotemporal, and perceptual information and is interpreted as having meaning (Kate Newell, 2021). In terms of media transformation, the key

lies in the change of material form. It is worth noting that the material dimension of media involves technical elements, which offer different expressive possibilities (Silvia Kurr, 2023). This transformation emphasizes that as painting images migrate to dynamic media, their visual information integrates with other sensory inputs such as sound and touch, creating a richer, more comprehensive sensory experience, which in turn enhances expressiveness and emotional communication. Combining visual perception theory with cross-media modality theory reveals the multi-layered mechanisms involved in the media transformation of painting images. This process not only optimizes the processing of visual information but also promotes the integration of multimodal information, significantly enhancing the audience's sense of immersion and emotional engagement.

In the visual expression of painting images, light, color, space, and motion are essential elements that contribute to the generation of visual meaning. Together, they participate in the process of image perception and aesthetic experience. In different media contexts, they are not presented in their original form but are mediated and re-represented through various media modalities. The material modality, through means such as screens, sound waves, and virtual technologies, processes and simulates space and perception, altering the material foundation of visual experience and revealing the fundamental role of media materials in meaning construction. The sensorial modality, through multisensory systems such as vision, hearing, and touch, presents the static qualities of light and color, as well as the changes of space and motion in the perceptual dimension. The guiding nature of the sensorial modality reflects the path selection of different sensory channels in information reception and cognition. The organizational nature of the spatiotemporal modality manifests in the structured presentation of information in time and space. The time modality introduces movement and rhythm through fixed or partially fixed sequences of time, giving the image dynamic and changing characteristics. The spatial modality, through different media interface features, creates spatial depth and illusion, constructing visual layers. The material modality, sensorial modality, and spatiotemporal modality form a framework for mediating the pre-symbolic process of interpretation. The synergistic interaction of these three media modalities allows visual information to be perceived, interpreted, and disseminated within the media system. This multimodal mediation mechanism model is illustrated below (Figure 1)
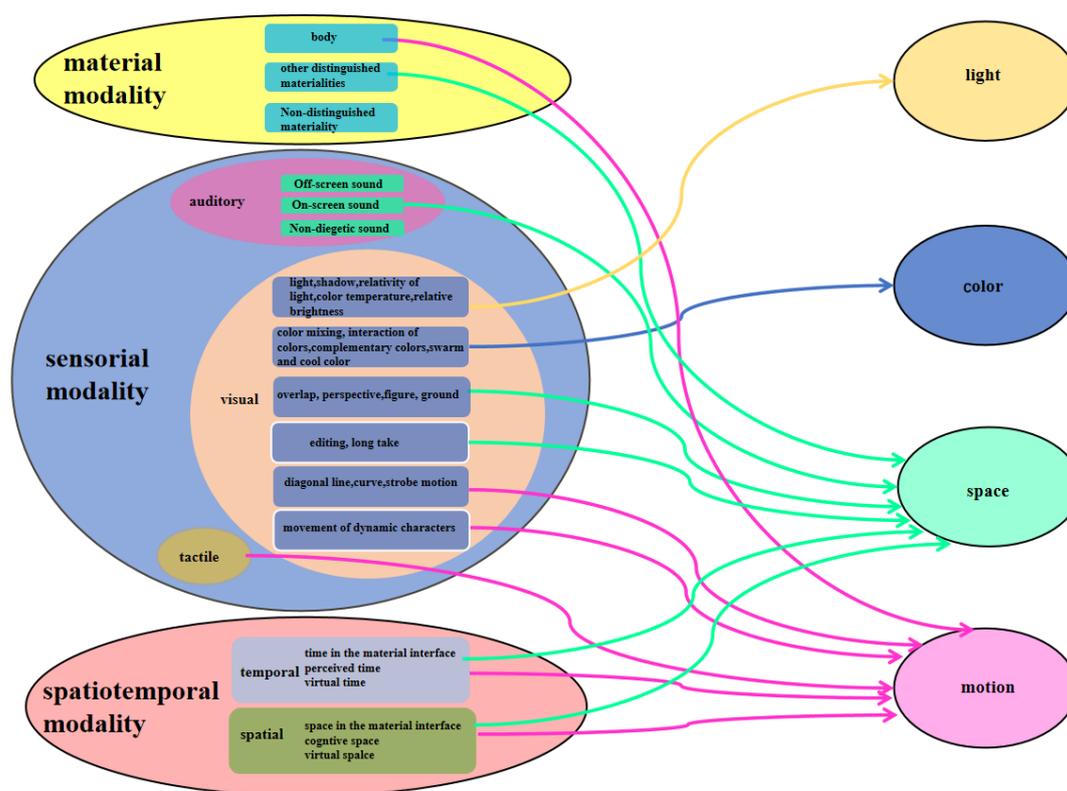
<Figure1> A Mediation Model of Media Modalities (Developed by the Researcher)

3. The Migration of Painting Images: Animation, VR, and Immersive Exhibitions

3.1 *Loving Vincent*: Expanding into Deep Space through Sensorial Modality

In the film, the material modality is manifested through the transformation of van Gogh's original paintings into dynamic images and continuously changing sounds. Given that the sizes of van Gogh's original paintings vary, while the aspect ratio of the film screen is fixed, many shots in the film required resizing of the original works. To meet the visual demands of the film, the creators had to appropriately expand or cut portions of the paintings to ensure a unified and coherent visual effect on the screen. The film primarily features works by van Gogh that depict all seasons, but the narrative time in the film is set around the late summer and early autumn before van Gogh's death. Therefore, the filmmakers needed to ensure temporal consistency throughout the film, adjusting the color, tone, and distinct seasonal features in works that did not align with the season. The post-production team then recorded the oil paintings frame by frame and digitally combined them into a unified oil painting film. This animation film, presented on a two-dimensional screen, on the one hand, mimics van Gogh's brushstrokes by placing paintings with a "van Gogh style" onto the screen. Through the use of perspective and depth of field, it dissolves the flat quality of the painting and creates a sense of three-dimensionality, presenting a spatial scene that approaches realism, thereby expressing spatiality

286

on the flat screen. On the other hand, the film breaks the limitation of relying solely on camera movements to dissolve the edges of the frame. Instead, it uses the elements within the painting as spatial media, imaginatively representing the transition from painting to film (Zeng,2021), thereby transforming the static into dynamic and expanding the visual scope of the scene.
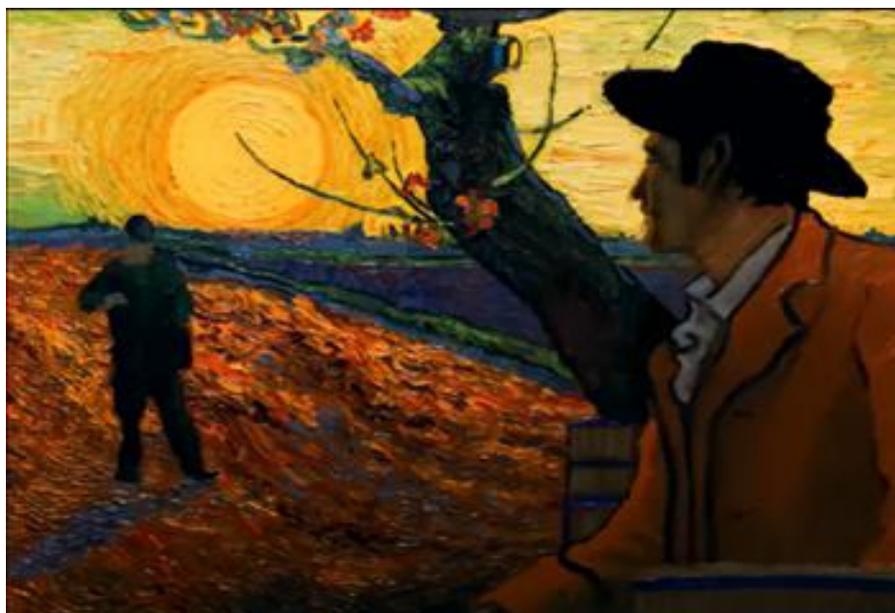
<Figure2> Complementary Colors—The Visual Transformation in *The Sower*

As shown in <Figure2>, In *The Sower,* the yellow sun contains a sufficient amount of golden tones, preventing visual conflict with the orange-yellow land. This allows the eye to connect the two colors through structural transformation, visually linking heaven and earth. The sower stands astride the boundary line, reduced to a flat silhouette of deep blue-black, forming a "positive-negative" relationship with the blazing background colors. The symbolic and repetitive sowing motion is rendered through mechanically bent arms at right angles, with seeds scattered in dense yellow dotted brushstrokes that trace a perceptual visual trajectory of movement. The physical motion of his body, however, remains continuous.

Within this visual mode, the sacred image of the laborer is conveyed while also highlighting the smallness and fragility of humankind before the forces of nature. The sower stands between heaven and earth, embodying the biblical "mediator between the heavens and the earth," echoing Van Gogh's letter in which he wrote of seeing a Christ-like eternity in the figure of the sower. This scene intertwines art, faith, and the suffering of life, transforming Van Gogh's painted imagery into an extension and expression of his spiritual belief and existential value.

In *Landscape with Carriage and Train* (1890) (see Figure 3), the visual composition reveals a carefully organized tension between space and time. In the foreground, the tilted carriage and dark brown soil contrast with the middle ground, where parallel furrows of wheat

fields stretch horizontally, while the background features a train and straight railway tracks cutting diagonally across the canvas at a 45-degree angle. This layered composition not only constructs a strong sense of depth but also introduces a temporal confrontation and symbolic interplay within the spatial arrangement.

The shape of the waving wheat can only be fully perceived from an aerial viewpoint, which captures its true overall contour. However, such a bird's-eye perspective makes it difficult to represent the carriage and distant train in a recognizable, conventional manner. To accurately depict these two moving objects, a side view would be the more ideal choice. By employing an Egyptian mode of representation, Van Gogh compresses the images of the carriage and the train into the pictorial surface, integrating information from different perspectives into a unified composition. In doing so, he breaks the traditional rules of perspective to achieve formal tension and narrative completeness on a two-dimensional plane, thereby creating a unique visual effect.

<Figure3> The Long Take Interpretation in the Animated Film *Loving Vincent*

In *Loving Vincent*, the carriage and the train are in motion, with direction and speed influencing their combination and separation (Arnheim,2019: 82). In the scene, the carriage moves to the left while the train moves to the right, creating a significant sense of separation. Under the long take, the subjective perception of speed differences is further intensified by the camera's movement, which enhances depth perception in the filmed scene. This is because when a moving camera sweeps across the scene, the perceived speed depends on the distance between the objects and the observer. The closer carriage appears to move much faster than the distant train. This sense of uniformity and difference in speed helps us determine distances and reinforces the clear spatial structure, effectively enhancing the contrast between the foreground and background.

Furthermore, the carriage and the train hold profound cultural symbolic meanings within

288

the symbolic mode. The carriage represents the rhythm of the old world, signifying an intimate connection between humans and nature, while the train symbolizes the speed of the new world, representing the rupture between industry and nature. These two carriers of time and civilizational logic coexist in the frame but never intersect. Through the visual organization of the scene, they are placed in the same frame, yet remain isolated from one another, unable to truly intersect. This visual fragmentation enhances the psychological sense of alienation, evoking an inner tension and sense of estrangement that the audience feels as they experience the shift of eras.

3.2 *The Night Café VR:* Exploring Media Space through the Material Modality

In VR animation, the user's bodily interface functions as a medium, participating in the perception and understanding of space through vision, touch, and proprioception. This mode of experience differs from the passive viewing of traditional animated films, as it is based on immersive perception and the active spatial cognition constructed through bodily presence.

When Van Gogh's symbolic motif, the sunflower, is placed on the piano in the left hall (see Figure 4), it appears as a visual object in the virtual space. The recognition of this object relies less on its shape itself than on the structural framework the shape creates. As Arnheim (2019) points out, three key frameworks influence the audience's perception of spatial orientation: retinal orientation, environmental orientation, and bodily movement. Bodily movement is determined by sensations from the muscles and the inner ear's balance mechanism (Arnheim,2019:102).

No matter how the position of our body, head, or eyes changes, we can always perceive the direction of gravity. In daily life, the sense of movement and the visual perception of environmental orientation are usually in harmony. For example, when we look up at the sunflower, the image may appear tilted because our head is raised and the viewing angle has changed. However, our body "knows" that the head is tilted; any movement of the eyes, head, or body is transmitted to the brain's motion perception center, which receives the information that the head is inclined. The feedback from these motion processes influences visual perception. Consequently, the brain automatically corrects the image, straightening it by attributing the movement to the head rather than the environment.

Through the interaction between visual perception of the tilted visual field and the proprioceptive awareness of one's tilted head, the viewer ultimately perceives the sunflower as standing upright. This intersensory coordination reflects how VR's embodied perception reinforces spatial realism and cognitive engagement in virtual environments.

<Figure4> Movement at the Bodily Interface 1

In visual perception, the meaning of "vertical" within the visual field differs from its definition in physics. When we stand upright, lie down, or tilt our heads, we can still perceive the objective or physical direction of gravity—this is known as environmental orientation. At the same time, however, we may perceive a flat book page or a painting lying on a table as having a distinct top and bottom. When we look down at a book on the table, the top of the page is not physically above our head but rather appears at the top of our visual field. This orientation, determined by the position of the image on the retina, is what Arnheim (2019) refers to as retinal orientation.

In the VR animation based on Van Gogh's works (as shown in Figure 5), the user follows Van Gogh as he rises from his chair, walks slowly toward the window, and looks up at *The Starry Night* outside. At this moment, the starry sky beyond the window seems to extend the physical space of the room, a visual continuity established through environmental orientation. On the floor, elements from another of Van Gogh's paintings, *Café Terrace at Night*, appear in the form of café chairs, thereby integrating visual motifs from multiple paintings into a single, continuous spatial scene.

<Figure5> Movement at the Bodily Interface 2

More importantly, in addition to the coordination between the retinal region and the visual environment, bodily movement also plays a crucial role in this process. When the user looks up at *The Starry Night*, the interaction between the visual perception of the tilted visual field and the proprioceptive sensation of the head's upward tilt allows the user to naturally perceive the buildings outside the window as standing upright before them, bathed in the vast starlit sky. Similarly, when the user looks down at the *Café Terrace at Night* scene on the ground, the collaboration between body and vision aligns the perception with the true direction of the ground, thereby enhancing the sense of immersion and realism within the scene.

When the user turns their head or leans their body in VR, the information that the head is moving causes the visual system to attribute this motion to the head itself, resulting in the perception that the head is moving while the surrounding environment remains still. This visual stability arises from our constant ability to perceive the direction of gravity, regardless of how the body, head, or eyes are positioned. Therefore, bodily proprioception enables the user to move—and even step into the space—and when the visual orientation conflicts with the environmental orientation, bodily proprioception functions as a coordinator to reconcile the two.

Ultimately, the interaction among these three mechanisms of spatial orientation not only explains the spatial relationships within the image but also instills in the user a sense of presence and motivation for exploration. The user is no longer a static observer but an active participant who steps into the pictorial space. Through the coordinated functioning of sensorial modalities and material modalities, the VR experience achieves a high degree of immersion.

In VR animation, the essence of the material modality lies in the intentional invocation and reinforcement of human spatial visual perception characteristics. The user's bodily movements can be fully recognized through motion detection, with head movement playing a crucial role in controlling the visual perspective. The creator uses Van Gogh's *Self-Portrait* as a template, rendering Van Gogh's image in three-dimensional form within the VR space, thereby transforming the original two-dimensional painting into an immersive virtual figure.

As shown in Figure 6, the sense of space within the VR environment arises from the light emitted by objects in the space entering the visual field. This input provides more abundant and powerful depth cues, and since no elements in the virtual space obstruct depth perception, the user's spatial awareness is significantly enhanced.
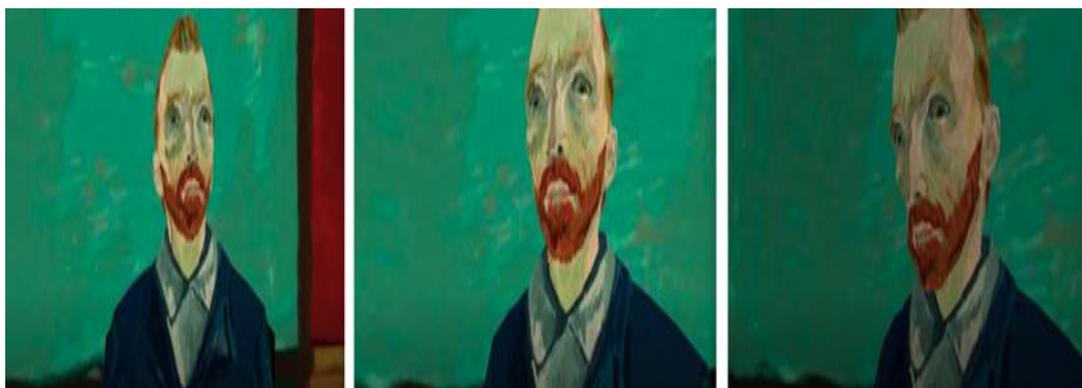
<Figure6> Spatial Expression of Visual Focus

This technology is not a two-dimensional reproduction of the original painting, but rather a reconstruction of the original light input. As the user moves within the VR space, the information generated by motion perception produces two depth cues: binocular disparity and accommodation. When the user records Van Gogh's image from different angles and distances, the two eyes must converge their lines of sight to integrate the views into a single stereoscopic image (Su, 2021).

Furthermore, the user can move within the virtual space using a controller—approaching or moving away from Van Gogh's figure. As Arnheim (2019) points out, when an object is close to the eyes, the angle formed by the visual axes is large; as the object moves farther away, the angle becomes smaller. To fuse the images seen by both eyes into a unified visual experience, the tension of the ocular muscles—responsible for eye movement and fixation— changes according to the object's distance from the eyes. This convergence movement itself serves as a perceptual cue for depth.

Similarly, the dynamic sensation from the ciliary muscles is used by the nervous system as an indirect indicator of distance. This focusing mechanism is activated through the gradient transition from blur to clarity within the visual field, guiding the user toward an accurate perception of depth and spatial structure in the virtual environment.

3.3 *The Starry Night by Van Gogh*:A Continuum from Temporal-Spatial Modalities to Imaginal Space

Given that temporal-spatial modalities are a subset of temporal-spatial medium modalities, and since the temporal-spatial medium modalities are composed of physical entities, all medium products possess temporal-spatial attributes that can be grasped by human cognition. In the time modality, the time displayed on the material interface cycles through fixed time sequences, forming a perceivable physical time. However, the viewer has the freedom to move and pause during the viewing process, thus creating a subjectively perceived time that varies across individuals. Meanwhile, virtual time is presented through different thematic spaces,

reflecting the artistic time of Van Gogh's creative career. This time is not linear historical narrative but a temporal structure constructed by the creator based on Van Gogh's artistic context, allowing viewers to sense the transformation of his emotions and style through spatial changes. In architectural exhibition spaces, communication and human experience cannot exist virtually as they do in literary works through language; instead, they truly unfold in the physical space displayed by the interface. The exhibition constructs a three-dimensional space that can be physically entered through the proportions of height, length, and depth. As the viewers move and pause within it, they form the process of generating cognitive space. In Van Gogh's immersive exhibition, different thematic spaces correspond to different periods of the artist's life, resulting in an overlapping of multiple virtual spaces.

In architectural exhibitions, space and time are inseparable. The existence of space is realized through the use of physical elements such as walls and floors, allowing the audience to experience a designed immersive experience within the real spatiotemporal context. As a temporal-spatial practice of a material medium, architectural exhibitions not only present works of art but also, through the tangible existence of space, carry the flow of time, enabling the audience's body, perception, and mind to actively participate in the generation of artistic meaning. Therefore, it is of critical significance to study the "human" as the medium linking the material space of architecture and the conceptual space of text. Whether it is the material interface space or the virtual space, the integration of the two requires the embodied perception and cognition of "the human" as its foundation. When exploring architectural modalities, two inherent concepts must be further considered: embodiment and perspectivity.

The content and form of contemporary art have become increasingly diverse. In particular, there is a growing number of works that require active audience participation. Hong Okjin (2015:37-64), from the visual dimension of perceptual aesthetic experience, conducted an in-depth analysis of the participatory characteristics of contemporary art aesthetics. He pointed out that through direct bodily involvement, the audience can immerse themselves in the artistic world constructed by the artist, while avoiding a sense of superiority in the audience's subjective awareness within their inner world. Therefore, the entire aesthetic experience can be understood as an interactive process. This view lays the theoretical foundation for further research on the interaction between architectural exhibitions and the human body. At the same time, the concept of embodiment proposed by Maurice Merleau-Ponty (1992:18) not only supports the inseparability of body and mind but also emphasizes perception in the world's sensory experience and representation. In immersive exhibitions, the audience enters the spatial domain formed by multidimensional virtual images through their own bodies as a medium. These digitally processed 720-degree panoramic images collectively create an immersive experience similar to a stage performance, transforming the act of viewing into a sensory

practice of bodily participation. In the field of performance studies, embodiment is often used to explore the stage experience of actors and the perception of the audience. Chiel Kattenbelt (2006) refers to "embodiment" as the effect of the actor's body on stage, considering the body's presence as a materiality within the theatrical space (Kattenbelt,M.J.,2006). This view resonates with Elleström's discussion of how the dancer's body simultaneously functions as both a technical medium and a medium product: "How can we distinguish the dancer from the dance?" (Lars Elleström, 2021). For architectural immersive exhibitions, the focus is on determining how the audience experiences the three-dimensional spatiality of the final medium product through kinesthetic perception. In other words, although the creation of the exhibition initially relies on the conceptual communication and cognitive transfer between the creator and the technical team, any exhibition, in order to transcend the physical realm of architecture and become a true work of art, it must ultimately be realized through bodily participation and experience.

<Figure7> Embodied Immersive Expression of Temporal Parallax in *Self-Portrait*

Van Gogh's *Self-Portrait* (Figure 7) exhibits distinct differences in brushstrokes, lighting, tones, and emotions. In immersive exhibitions, these self-portraits are often juxtaposed, overlapped, and dynamically switched, forming a spatiotemporal continuum of the self. As viewers move from one projection surface to another, they may see different self-images of Van Gogh from various periods; the interplay of light and shadow, along with dynamic projection techniques, creates a flowing and overlapping effect between these faces. The viewers' bodily movement and the changing images intertwine in time, creating a temporal parallax. Although the space displayed on the material interface, presented through a panoramic screen and multi-projection, remains a flat dynamic image, the viewers receive different visual information as they move within the exhibition hall. As Arnheim pointed out, the human eye projects optical images as two-dimensional pictures, and the brain cannot fully grasp the three-

dimensional nature of an object from a single perspective. Temporal parallax arises from the positional shift of the observer, causing differences in visual images. These varying visual forms are psychologically synthesized into a unified three-dimensional image, thereby assisting the viewer in mentally constructing a depth sense that transcends the flatness of the image and enhancing the immersive experience. This phenomenon of depth constructed by the body is at the core of embodied cognition. Immersion does not arise from the space itself possessing depth but from the interaction between the body and the space, which allows me to perceive that this space has depth significance for me.

Arnheim's dynamic mode of observation resonates with the rhetorical technique of "periegesis" used in ancient descriptions of places, landscapes, and architecture, as well as with the camera techniques in film. Even when the retinal perspective mechanism causes distortions or visual illusions of objects, the human mind can still perceive three-dimensional objects as a whole by organizing, completing, and synthesizing their structures. This depth of real space gained through bodily movement differs from the three-dimensional effects generated by visual mechanisms in VR films. The depth in VR animation comes from the fusion of two images captured by the two eyes. The depth cues extracted by the brain are based on the differences between these two images, a phenomenon known as binocular spatial parallax. In the experience of architectural exhibitions, viewers actively move and perceive the space with their bodies, directly experiencing the depth of the space. This embodied perception provides a more immediate immersive effect. The cognition of architectural space-time eternity, the uniqueness of the human eye and mind relationship, and the media differences shaped by various historical stages and regional characteristics collectively determine the artistic and aesthetic attributes of the exhibition.

In summary, embodiment reveals the experiential mechanism in immersive exhibitions, where the viewer uses their body as a medium of perception and action as a mode of cognition. The viewer's body is not only the subject of perception but also becomes an active participant and constituent in the generation of artistic meaning. However, embodied experience does not exist in isolation; it is always intertwined with visual perception. As the history of perspective in art has shown, the way we observe determines the relationship between the subject and the world. Therefore, when further exploring the spatiotemporal structure of architectural exhibitions, attention must also be given to "perspectivity"—that is, how the viewer, through the organization of vision and the construction of space, establishes the relationship between the self and the world.

The origin of the term "perspective" in art history and the history of science can be traced back to Arabic geometry (i.e.,mathematics) or Western theories of perception (i.e., aesthetic theory). In painting, perspective creates dynamic visual relationships that allow the viewer to

inhabit spaces they would never physically reach. In other words, presence and absence are interwoven within this inseparable relationship. With the publication of Edwin Panofsky's *Perspective as Symbolic Form*, the concept of perspective expanded from geometric measurement to cultural and symbolic systems. Elkins further noted that "the modern concept of perspective has a schizophrenic nature, as it encompasses two irreconcilable aspects": one is "the mathematically rigorous branch of the discipline, from Brunelleschi's experiments to the latest software systems of perspective," which represents the contemporary understanding of perspective in the field of architecture; the other "includes a vast range of meanings, from subjectivity to eternity, extending across almost all areas of philosophy, literature, and political discourse." This latter dimension is referred to as metaphorical perspective, which Elkins explains as "our perspective—it is the perspective through which we think, as well as the one that describes how we observe the world and construct ourselves as subjects of observation" (Elkins & James, 1994). This theory of the subjectivity of viewing is newly embodied in contemporary immersive art exhibitions. In the study by Görgülü, Aslı Çekmiş, and İlgi Hacıhasanoğlu (2012), a comparison was made between the immersive Van Gogh exhibition and traditional museum exhibitions. The immersive Van Gogh exhibit breaks the stereotype of viewing masterpieces from a distance in a silent art gallery. Viewers are free to walk around or sit on the floor in the gallery, interacting with the artwork at will, thus revolutionizing the way art is experienced and awakening sensory impact. This display model, using large-scale canvases as a medium, provides the audience with a completely new art experience.

In the immersive exhibition *Van Gogh:Starry Night* (Figure 8), the presentation of *Starry Night* is no longer a simple digital reproduction of the original work. Instead, it undergoes a creative reconstruction through the displacement, transition, and movement of images and elements. In traditional painting, the spatial perspective of *Starry Night* is still based on the "mathematical perspective" principles established since Brunelleschi, where the village in the foreground, the hills in the middle ground, and the swirling starry sky in the background all form a stable viewpoint at the center, with the world organized within a rational geometric system. However, when the work is transferred to an immersive projection environment, this perspective is completely broken and reconstructed. The projection walls, floors, and even the audience's bodies become carriers of the image, causing the vanishing point of the perspective to spread across the entire space. As a result, the image no longer originates from a single viewpoint but rather regenerates spatial relationships as the viewer moves through the space.

<Figure8> Representation of Perspective in the Spatial Viewpoint of *Starry Night*

Perspective, therefore, is not only a mathematical tool for measuring the world but also a metaphor for how we view and construct the self. In the immersive *Starry Night*, this metaphorical perspective is dramatically intensified. The audience is surrounded by vast rotating nebulae and flowing light, and the external viewing point provided by traditional paintings is completely dissolved. Viewing is no longer about me watching the painting, but rather about being engulfed by it. Perspective is transformed from an external geometric order into an internal, subjective psychological space experience that is enveloped, absorbed, and even disintegrated. In other words, in the immersive experience, the viewer becomes the material of perspective, rather than its center. This experience reveals what Elkins refers to as the schizophrenic nature of perspective. In immersive *Starry Night*, the precision of mathematical perspective coexists and conflicts with the emotionality of metaphorical perspective, creating a schizophrenic landscape that is characteristic of modern viewing experiences. This phenomenon can also be confirmed through Benjamin's theory of the aura. With the advent of the age of technological reproduction, the aura of art works gradually fades. The work is stripped from its original context and gains a new, mass-oriented form of existence. Therefore, some researchers have pointed out that this type of exhibition model aligns with the so-called "experience economy"demand, where businesses thrive by providing unforgettable consumer experiences, and its core has transcended the simple exchange of material goods (Mondloch&Kate,2022). In such an environment, the audience becomes the consumer of the experience, and is also a crucial part of the art reproduction and reproduction mechanism.

It is evident that perspective in immersive exhibitions undergoes a transformation from a geometric paradigm to a perceptual metaphor. The rational logic of mathematical perspective and the emotional dimension of metaphorical perspective intertwine, extend, and ultimately manifest in the embodied perception of the viewer within the spatial context of digital media. The immersive *Starry Night* reconstructs the viewing mechanism through technological means,

transforming perspective from a structural representation into a perceptual experience, thereby revealing the logic of the reconfiguration of the subject-space relationship in contemporary visual culture.

Conclusion

Through a comparative exploration of the common characteristics between Van Gogh's works and *Loving Vincent*, *The Night Café VR*, and *Van Gogh: Starry Night*, this paper develops a relevant but seldom-practiced analytical approach. The potential for continuous creation of new cultural objects and practices in the field of painted image art, fueled by advancements in modern technology, paves the way for a possible development path.

In fact, these projects share a common feature: they all break the boundaries of the canvas and place significant emphasis on the concept of "experience," which is closely related to bodily perception and sensory stimulation, expanding the understanding and accessibility of Van Gogh's works. Through a broader and more detailed analysis of the audience responses and interactive effects stimulated by these works, it becomes clear that the material characteristics of these objects provoke different reactions. *Loving Vincent* offers a highly aesthetic cinematic experience that reshapes our understanding of audiovisual art through the reconfiguration of sensory modalities. In contrast, *Van Gogh:Starry Night* and *The Night Café* use digital technology to create more participatory and spatially immersive experiences. *The Night Café* utilizes material modalities to overlay a layer of digital information onto the perceptual layer of the real world during the experience, immersing the audience in a completely new media space. The *Van Gogh: Starry Night* immersive exhibition, on the other hand, relies on real physical space, combining film and panoramic views to enhance the sense of virtual illusion in a concrete setting, creating an art experience that blends reality and imagination.

Therefore, creators use multimodal approaches to media, generating a strong immersive experience in the fissures and fault lines of sensory modalities. First, there is the continuous extension of the boundaries of performing arts; second, the structural connections hidden between seemingly heterogeneous media; and third, the aesthetic narrative potential sparked by media transformation and adaptation. Cross-media creation is not merely the layering of multiple media; it is a dialogue and release of tension between media, encouraging the audience to shift between viewing, listening, and interacting, thereby reconstructing meaning in fragmented sensory experiences and forming multidimensional perceptual participation.

From the perspective of media modalities, *Loving Vincent*, *Van Gogh: Starry Night*, and *The Night Café* expand the boundaries of artistic experience through the reconstruction of material, sensorial, and spatiotemporal elements as pre-symbolic layers. However, there are

also certain limitations. The sensorial modality remains predominantly visual, which causes the immersive experience to become somewhat one-dimensional in terms of perception. The artistic experiences created through digital media do not fully encompass all aspects of the painting image. Interactive design is often based on technical presets, and the audience's participation remains at the superficial level of interaction. Future development should enhance the integration of multimodal elements, combining artificial intelligence and sensing technologies to create dynamic interactivity in the perceptual layer. At the same time, it should promote the return of material modalities, rebuilding the bodily and live presence of art. Humans always center their understanding and definition of the world around their own circumstances and cognitive systems. Based on this, immersive art should shift from a technical display to a media reflection, making the transition from aesthetics to sensorial studies.

· Declaration of conflicting interest: No conflicting interest.

·Funding statement：No Funding support

·Ethical approval and informed consent statements：There are no ethical issues, and informed consent from the first author (in oral form) has been obtained.

·Data availability statement：Data can be used indefinitely.

·Any other identifying information related to the authors and/or their institutions, funders, approval committees, etc, that might compromise anonymity：No.

COI missing: NO.

References

Arnheim,R.(2019).*Art and visual perception*(S.Teng &J.Zhu,Trans.).Sichuan People's Publishing House.

Baranseli,E.S.(2024).The new aura of the artwork in the digital age of infinite reproduction: A study on the Van Gogh Alive exhibition.*The Journal of Academic Social Science*, 67(67), 196-211.

Elkins,J.(1994).*The poetics of perspective*. Cornell University Press.

Elleström,L.(2021).*Beyond media borders*(Vol.2):*Intermedial relations among multimodal media*.Palgrave Macmillan.

Görgülü, A.Ç.,&Hacıhasanoğlu İ. (2012, November). A new spatial experience of art via technology. *In ICONARCH International Congress of Architecture and Planning,*244–253.

Guo,W (2025). A study on the construction of POI sequences in VR animation media narratives: Focusing on *The Night Café: VR Van Gogh*. *Journal of the Korean Society of Contents*, 25(6), 272-284.

Hong,O.(2015).The possibility of a phenomenological approach to aesthetic experience. *Art and Media*,14(1),37–64.

Kurr,S.(2023).Towards ecological thinking through modal changes:Weird media transformation in Lars von Trier's film Melancholia. *Ekphrasis: Images, Cinema, Theory, Media,* 29(1), 12–24.

Lee,S.(2024) Media as embodied cognitive systems in the posthuman era:Media ecology and embodied cognition perspectives.*Journal of Digital Media Studies*, 21(2), 63–88.

Lee,S.(2024).A study on "defamiliarization" in experimental animation. *Animation Studies*, 20(3), 50-79.

Merleau-Ponty,M.(2001).*Phenomenology of perception* (Z.Jiang,Trans.).The Commercial Press.

Mondloch,K.(2022).Theinfluencers:Van Gogh immersive experiences and the attention-experience economy. *Arts*, 11(5), 90.

Newell,K.(2021).*Transferring handmaids: Iconography, adaptation, and intermediality*. In L. Elleström (Ed.), Beyond media borders (Vol. 2, Intermedial relations among multimodal media, p. 201). Palgrave Macmillan.

Sainae,&Jang,S.(2021). A discussion on the existence of painting from the perspective of media convergence: Based on the analysis of "Immersive Experience of Van Gogh's Art." *Journal of the Korean Society for Science and Art Convergence*, 39(1), 165-180.

Su,X.(2021).On the unity of body and mind in VR painting space: A study of spatial phenomenology from the perspective of Maurice Merleau-Ponty.*Theoretical Studies in Literature and Art*, 41(5),144–156.

Tian, L., &Zhang,Y.(2019).The innovation of Greimas' structuralist narratology in *Loving Vincent*. *New Films*, 72-75.

Wang,N.(2018).On VR film scenario narrative based on embodied perspective.*Contemporary Cinema*, 111-115.

Yao,T.(2018).The convergence of oil painting and film creation: An analysis of pastoral imagery in *Loving Vincent*. *Film Review*, (15), 101–103.

Zeng,J.(2021).Media technology imagination: A possible art theory.*Comparative Literature in China*, (1), 12.