

Knowledge Graph for Intelligent Generation of Artistic Image Creation: A New Annotation System

Authors: Kaixin Jia, Kewen Zhu, Deng Huanghuang, Qiu Yiwu, Ding Shiyong, Ding Chenyang, Li Zejian, Huanghuang Deng

Date: 2025-11-06T00:00:00+00:00

Abstract

This study aims to construct a unified, systematic, and referential knowledge framework for art image dataset annotation to address the issues of ambiguous definitions and inconsistent results arising from the lack of common standards during the annotation process. To achieve this objective, based on the compositional principles of art images and incorporating the “Structured Theory of Visual Knowledge” proposed in *On Visual Knowledge*—which posits that visual knowledge requires precise expression of spatial shapes and dynamic relationships through “typical-category” and “hierarchical structures”—this study resonates with the “Dual Attributes of Cultural Elements” theory in *Cultural Composition* to comprehensively examine both Chinese and Western art theories, and innovatively integrates a Chinese cultural perspective to construct a hierarchical and systematic knowledge graph for art images. This graph deconstructs the core visual language of art images and references and supplements Western art theory with the unique spatial theory and symbolic system of Chinese painting. By transforming qualitative artistic concepts into a clear structured framework, it not only conforms to the human cognitive principle of “visual knowledge preceding verbal knowledge,” but also provides an explainable and inferable visual knowledge foundation for AI art image generation and cross-cultural art analysis, ensuring high quality and consistency of annotation data, and offering critical support for art intelligence research in the AI 2.0 era.

Full Text

Preamble

Journal of Zhejiang University (Engineering Science)

Knowledge Graph for Intelligent Generation of Artistic Image Creation: Constructing a New Annotation Hierarchy

JIA Kaixin¹, ZHU Kewen², DENG Huanghuang^{3*}, QIU Yiwu⁴, DING Shiyang⁴, DING Chenyang⁴, LI Zejian¹

(1. College of Software Engineering, Zhejiang University, Ningbo 315199, China; 2. College of Art and Archaeology, Zhejiang University, Hangzhou 310028, China; 3. College of Computer Science and Technology, Zhejiang University, Hangzhou 310028, China; 4. Hangzhou Zaowu Cloud Technology Co., Ltd., Hangzhou 310028, China)

*Corresponding Author

Abstract: Our study aims to establish a unified, systematic, and referable knowledge framework for the annotation of art image datasets, addressing issues of ambiguous definitions and inconsistent results caused by the lack of common standards during the annotation process. To achieve this goal, a hierarchical and systematic art image knowledge graph was constructed. It was developed based on the composition principles of art images, incorporating the “structured theory of visual knowledge” proposed by Academician Pan Yunhe in *On Visual Knowledge*—which states that visual knowledge must achieve precise expression of spatial forms and dynamic relationships through “prototype-category” and “hierarchical structure.” Through in-depth review of Chinese and Western art theories and pioneering integration of the Chinese cultural perspective, this graph took shape. The core visual language of art images was deconstructed by this knowledge graph. Meanwhile, the unique spatial theory and symbolic system of Chinese painting were compared with and supplemented by Western art theories. This graph converts qualitative artistic concepts into a clear structured framework. It not only conforms to the cognitive law that “visual knowledge takes precedence over verbal knowledge” in humans but also provides an interpretable and inferential visual knowledge foundation for AI art generation and cross-cultural art analysis. It ensures the high quality and consistency of annotated data, thus offering key support for art intelligence research in the AI 2.0 era.

Key words: Art Image; Knowledge Graph; Chinese Culture; Painting Language; Structured Framework

This research stands at the intersection of two major contemporary trends. On one hand, artificial intelligence technology is advancing toward the AI 2.0 era, characterized by the co-driving of data and knowledge, with particular emphasis on visual and multimodal intelligence. On the other hand, the excavation of “cultural genes” and the digital preservation of cultural heritage have become crucial national strategic imperatives for enhancing cultural confidence and promoting civilization inheritance. However, in the specific domain of artistic creation, current AI technologies represented by deep learning still face fundamental limi-

tations. While deep neural networks have achieved remarkable progress in tasks such as artwork style classification and artist identification, they remain essentially “black box” systems based on statistical pattern matching. They can recognize “what” something is (e.g., this is a Baroque painting) but cannot explain “why” it is so. These models cannot comprehend that a painting exhibits a particular style due to the intrinsic combinational logic of its composition, lighting, color, brushwork, and other visual language elements, nor can they perceive the cultural concepts and aesthetic paradigms deeply embedded in these visual choices. This phenomenon represents a “semantic gap” that urgently needs to be bridged in computational art.

To address this challenge, this study constructs the “Painting Language Knowledge Graph” (PL-KG). This knowledge graph is not merely a simple image tagging system but a structured, interpretable knowledge ontology that deconstructs the creative principles, technical theories, and aesthetic concepts accumulated over centuries of art history into a hierarchical, computable, machine-readable language. The core argument of this paper is that PL-KG provides an effective and innovative framework for the representation and creation of artistic expression through the concrete, operational implementation of two core theories: “Visual Knowledge” [?] and “Cultural Constitution” [?]. Its primary and fundamental innovation lies in its systematic, non-hierarchical integration of Chinese and Western visual paradigms. Rather than treating concepts from Chinese painting theory such as “scattered perspective” and “negative space” as supplements or exceptions to the Western linear perspective system, PL-KG positions them as equally important, thereby constructing a more universal and inclusive global visual grammar. This report systematically elaborates on the theoretical foundations, structural design, and profound value of PL-KG in advancing cultural digitization and intelligent generation of artistic image creation.

1 Theoretical Foundations: From Cultural Default to Visual Knowledge

The philosophical and theoretical cornerstone of this knowledge graph construction derives from Academician Pan Yunhe’ s profound reflections on modern design and artificial intelligence, which combine his “Cultural Constitution” theory with the “Visual Knowledge” framework. This theoretical integration not only provides justification for PL-KG but also endows it with deep cultural connotations and a cognitive science foundation, fundamentally elevating its design philosophy beyond purely formalist analysis.

1.1 Critique of “Cultural Default” and Cultural Constitution Theory

The starting point of this research is a critique of the pervasive phenomenon of “Cultural Default” in modern design, a critique reflected in Academician Pan Yunhe’ s reflections on Bauhaus design philosophy. As the cradle of modernist

design, Bauhaus explored abstract laws of art and design, advocating the reduction of works to basic geometric forms such as cubes, spheres, and cones. Due to its simplicity, rationality, and adaptability to industrial materials, this approach gradually became the mainstream after the 1950s. However, when this geometric formalism proliferated as the “default option” in design, its inherent defects became glaringly apparent: monotony, homogeneity, and a “lack of human touch” resulting from the stripping away of specific cultural contexts. The root of this phenomenon is identified as the “cultural default in design” [?].

To counteract “Cultural Default,” the “Cultural Constitution” theory was proposed [?]. The core of this theory is first to identify and cherish “Cultural Elements”—the most distinctive components of a culture that carry its identity, such as the unique roof and beam structures of Chinese architecture, Chinese characters and calligraphy, and the facial makeup and costumes in opera. These are the “branches and leaves” of culture, the fundamental embodiment of cultural uniqueness. Subsequently, “Cultural Constitution” is defined as a design methodology—the conscious and systematic application of these cultural elements in innovation. It transcends simple geometric formal composition, advocating that design should be integrated with profound cultural matrices to achieve aesthetic richness and emotional resonance. PL-KG represents the extension and application of this idea in the analytical domain, aiming to analyze the cultural constitution in existing artworks. By explicitly listing “Chinese characters and calligraphy,” “opera and costumes,” and “architecture” as cultural elements, and by juxtaposing unique visual laws from Chinese and Western art (such as “scattered perspective” and “linear perspective”), PL-KG elevates cultural specificity from a vague background factor to a core dimension of analysis, thereby systematically addressing the issue of “Cultural Default.”

1.2 Visual Knowledge Framework

If the “Cultural Constitution” theory infuses PL-KG with a humanistic soul, then the “Visual Knowledge” framework constructs its cognitive science and artificial intelligence skeleton. This theory posits that human visual memory differs from verbal memory—it is a form of knowledge that can be rotated, folded, scanned, and analogized, possessing unique spatiotemporal attributes [?]. Traditional symbolic knowledge representation (such as rules and semantic networks) struggles to capture the essence of this imagistic thinking. The design of PL-KG represents a comprehensive operational implementation of the “Visual Knowledge” theoretical system, precisely mapping its hierarchical structure onto the three core elements of visual knowledge: visual concepts, visual propositions, and visual narratives.

1.2.1 Visual Concepts Visual concepts are the basic units of visual knowledge with complex internal structures. Each knowledge node in PL-KG is designed as a structured “visual concept.” A visual concept consists of “prototype” and “category” [?]. In PL-KG, a specific knowledge node such as “S-shaped

composition” serves as the “prototype,” while its detailed definition (applicable scenarios: rivers/mountain paths/curved human dynamics; core function: enhancing depth) and quantitative standards (quantitative specification: S-shape occupying 60% of picture length) jointly define the “category” of this prototype. Through this “prototype + category” structure, PL-KG successfully transforms a originally qualitative and ambiguous artistic term into a clearly bounded, formal entity that can be understood and applied by computers.

Visual concepts themselves contain sub-concepts, presenting a hierarchical organizational relationship [?]. PL-KG’s multi-level directory system is a direct manifestation of this idea. For example, the specific visual concept “S-shaped composition” belongs to its upper-level concept “geometric composition,” which in turn belongs to “composition type,” ultimately falling under the top-level dimension “composition generation.” This top-down tree structure not only conforms to human cognitive habits but also constructs clear reasoning paths for visual knowledge.

Some visual concepts should also include elements composed of typical actions [?], which in PL-KG are embodied in descriptions of dynamic elements in artistic images, such as “brushwork rhythm” under “urgent-slow rhythm” (rapid brushwork = excited emotion, gentle brushwork = calm emotion), and tags like “trembling brushwork” and “rotating brushwork” in “calligraphic brushwork.” These tags go beyond describing the final form of the painting to capture the dynamic trends and emotional tensions inherent in the creative process.

1.2.2 Visual Propositions Visual propositions describe the spatiotemporal relationships of visual concepts [?]. Accordingly, annotations within the PL-KG framework can describe the abstract aesthetic characteristics of artistic images as multiple visual concepts and their interrelationships. The spatial relationships of visual concepts are presented through scene structures, which describe positional, distance, internal/external, up/down, left/right, and front/back relationships between objects [?]. The scene structure of an artwork can be described through combinations of tags from dimensions such as composition, space, and color in PL-KG. Taking Leonardo da Vinci’s *The Last Supper* as an example, its scene structure can be represented as a collection of visual concepts including “one-point perspective,” “focal composition,” “relative symmetry,” “horizontal movement line,” and multiple “triangular compositions” formed by figure groupings. These tags collectively explain the painting’s stable, solemn, and dramatically focused spatial layout.

Visual concepts’ temporal relationships are presented through dynamic structures, describing changes such as evolution, displacement, and action of different objects. In PL-KG, this is manifested in tags describing rhythm, cadence, emotional atmosphere, and brushwork dynamics.

1.2.3 Visual Narrative Visual narrative consists of an ordered set of visual propositions [?]. For instance, PL-KG’s definition of “scroll narrative per-

spective” provides a basis for generating strongly narrative artistic images. The viewing method of “changing scenery with each step” in many ancient scroll paintings is essentially a continuous visual narrative, where each pictorial segment can be analyzed as an independent visual proposition, and the ordered combination of these propositions constitutes a complete narrative.

Thus, PL-KG is not constructed from scratch but integrates the “Cultural Constitution” theory from the field of humanistic design with the “Visual Knowledge” framework based on cognitive science and artificial intelligence technology. This theoretical unification not only demonstrates the internal rationality of the knowledge graph but also elevates it beyond a general technical tool to become an epistemological bridge connecting humanities and science.

2 PL-KG: A Structured Machine-Readable Framework

Current art knowledge graphs primarily serve artists, art students, and the general public for creative guidance, lacking structured art knowledge systems specifically designed for artificial intelligence and machine readability. To fill this gap, this study systematically reviews authoritative works in Chinese and Western painting art, extracting from 20 classic theoretical works, two art dictionaries, and over 30 journal papers to develop PL-KG encompassing the following seven dimensions: composition generation, shape and form generation, perspective and space generation, light and shadow generation, color relationship generation, brushwork and texture generation, and edge relationship generation.

2.1 Labels for Artistic Image Composition Logic

Composition is the “backbone of a painting” [?]-it organizes visual elements, guides the viewer’s gaze, and establishes the structural foundation for the work’s theme and emotion. PL-KG deconstructs composition into multiple sub-dimensions including “composition core objective,” “composition type,” “picture fullness,” “viewpoint structure,” “visual guidance,” “visual balance,” and “rhythm and cadence.” The “composition core objective” reflects the functional orientation of composition, echoing Hu Wenxi et al.’s discussion on the relationship between composition and emotional expression [?]. The classification of “composition type” synthesizes Sattarov Farxod’s research on symmetry [?], Kaidong W’s systematic exposition on geometric order [?], and Roberts’ summary of various picture structures [?]. The establishment of “picture fullness” and “viewpoint structure” simultaneously covers Western “focal composition” based on single-point perspective and Chinese painting’s unique “scattered-point composition” [?] and “negative space” aesthetics [?]. The settings for “visual guidance,” “visual balance,” and “rhythm and cadence” are based on Arnheim’s [?] and Roberts’ [?] theories regarding eye movement flow, shape and color repetition, and other visual principles, forming the hierarchical vocabulary of this knowledge graph.

To demonstrate the explanatory power and applicability of this classification

method, in the Western art domain we select Leonardo da Vinci's *The Last Supper* as an example—a masterpiece that epitomizes Western classical composition principles. In terms of viewpoint and space, the work employs strict one-point perspective (also called parallel perspective), where all architectural lines converge at Jesus's head, making him undeniably the visual center of the picture [?]. This focal composition design highlights Jesus's central role in the narrative through perspective guidance [?]. Regarding balance and structure, the picture presents a highly symmetrical composition, particularly in the background architecture and figure groupings (two groups of three on each side of Jesus), demonstrating rigorous relative symmetry. Jesus himself forms a stable equilateral triangle, while the disciples' gestures expressing shock or surprise create a harmony between stability and dynamism within the triangular composition constraints, showcasing da Vinci's masterful control over picture layout [?]. The table's horizontal lines constitute a strong horizontal movement line, imparting a calm and solemn tone to the picture. In terms of thematic delivery, the entire composition serves a clear narrative guidance objective: capturing the disciples' dramatic reactions the moment Jesus says, "One of you will betray me."

In the Chinese art domain, we examine Fan Kuan's *Travelers among Mountains and Streams* as a classic work exemplifying Chinese compositional philosophy. Regarding viewpoint and picture fullness, the work assigns the foreground, middle ground, and background a progressive ratio of 1:3:9, "leaving the largest portion of the picture for the distant main peak" [?]. This monumental mountain layout creates a "critical visual leap," representing a classic example of the "high distance" composition from Guo Xi's "Three Distances" method described in *Lofty Ambitions of Forests and Springs*: "Viewing the mountain peak from below is called high distance" [?]. In terms of balance and guidance, the picture achieves visual size balance between the massive main mountain and the small path in the foreground. The towering peak is rendered in heavy ink, while a bright waterfall cascades from beside the main peak. The viewer's gaze is ultimately guided to a column of tiny travelers on the winding road in the foreground [?], forming a vertical movement line from top to bottom that reinforces the sublime grandeur of the mountain. In thematic delivery, unlike *The Last Supper*'s narrative-serving composition, *Travelers among Mountains and Streams* places the massive mountain at the picture's center for lyrical emotional reinforcement—through the size contrast between humans and nature, it conveys political implications and the Daoist "unity of heaven and humanity" concept of reverence for nature [?]. The above content can be described by the vocabulary proposed in this knowledge graph, forming structured annotations to assist in training and inference applications for text-to-image large models.

2.2 Labels for Artistic Image Shape and Form

This dimension focuses on how artists create three-dimensional, richly textured, and subjectively expressive images on a two-dimensional plane. It encompasses

“volume shaping,” “shape treatment,” “shape type,” and “positive and negative form” (figure-ground relationship). Among these, the classification of “volume shaping” is based on Arnheim’s visual perception theory [?], Johannes Itten’s color theory [?], and artistic anatomy [?, ?]. The classification of “shape type” builds upon Cézanne’s geometric generalization concept [?] and the “picturesque” aesthetic’s advocacy for natural forms [?]. The three strategies of “shape treatment” have clear artistic movement support: deformation and exaggeration correspond to Expressionism [?], simplification and generalization inherit from Cézanne’s modernist concepts [?, ?], and deconstruction and reorganization derive from Cubist creative methods [?, ?, ?, ?]. The classification of “positive and negative form” is established on Arnheim’s figure-ground relationship theory and its verification in computational vision [?, ?].

Using Leonardo da Vinci’s *The Last Supper* as an example: in terms of structure and volume, each disciple’s body demonstrates the artist’s profound understanding of human anatomy, with solid volume achieved through light-dark transition from light to dark. In shape treatment, da Vinci conveys the disciples’ reactions to the news through varied postures—these exaggerated and expressive gestures and body twists make the picture highly dynamic: James the Greater spreads his arms wide with astonishment; Peter leans forward emotionally; Philip points both hands to his chest, expressing innocence and earnestness [?].

Using Fan Kuan’s treatment of mountain and rock forms in *Travelers among Mountains and Streams* as an example: in terms of texture expression within volume shaping, Fan Kuan’s creation of “raindrop texture stroke” — “using the brush from squatting to slanting upward with rapid exit, interspersed with dense clusters of raindrop-like dots” [?]-effectively shapes the volume of yin-yang and the hard surface texture of northern mountain rocks. In conveying natural organic shapes, compared to mountains, rocks, and forests, Fan Kuan’s lines depicting figures and animals are more flexible and relaxed, thereby expressing the vitality of the subjects [?]. In terms of positive and negative form, the majestic, solid, heavily inked mountain body constitutes the positive form of the picture, while the negative space cleverly used around it suggests flowing clouds and misty atmosphere, undoubtedly blending virtual and real harmony into the overall stable heaviness [?]. It is evident that the vocabulary of this knowledge graph can describe most of the above appreciation content.

2.3 Labels for Perspective and Space Generation in Artistic Images

Spatial representation is a core issue in painting. PL-KG’s greatest innovation in this dimension is the juxtaposition of Western and Chinese perspective systems and the proposal of integrated application possibilities, thereby providing a unified analytical framework for these two different spatial philosophies.

The Western linear perspective system was systematized during the Renaissance by Alberti and others, based on geometric and optical principles aimed at creating a scientific, precise three-dimensional illusionistic space on a two-dimensional

plane through mathematical calculation. In contrast, the Chinese painting spatial system focuses more on expressing the artistic conception perceived by the mind, unconstrained by fixed viewpoints, and pursuing an experiential space that is “accessible and inhabitable.” PL-KG not only defines the core concepts of these two systems separately—such as Western one-point perspective, two-point perspective, and Chinese scattered perspective, Three Distances method—but more importantly, through the forward-looking label of “Chinese-Western perspective fusion application,” it provides tools for analyzing modern artworks that integrate both systems.

systematically compares the differences between the two perspective systems.

Table 1 Comparison of Perspective Systems

Western Linear Perspective	Chinese Conceptual Perspective
Theoretical Basis	Euclidean geometry and optics
Viewpoint	Fixed single viewpoint (“visual cone”)
Spatial Goal	Create rational, measurable, illusionistic 3D space
Time Expression	Capture single, frozen moment (“snapshot”)
Key Concepts	Vanishing point, horizon line

2.4 Labels for Light and Shadow Generation in Artistic Images

Light and shadow are core languages in painting for shaping form, guiding narrative, and creating atmosphere. PL-KG deconstructs the use of light and shadow into three hierarchical levels—“light and shadow function,” “light source and rays,” and “light-dark system”—forming a complete logical chain from “why use light” to “where light comes from” to “how to light.”

In terms of light and shadow function, da Vinci uses light to guide the viewer’s gaze to the picture’s core: Jesus sits in the center, his face illuminated by a bright window behind him, which not only shapes the figure’s three-dimensionality but also expresses the implication of Jesus as the “Light of the World.” From this point onward, light in painting has carried both symbolic meaning and spatial creation functions [?]. In the light-dark system, the traitor Judas’ s face is mostly obscured in shadow, contrasting with the bright Jesus. His body leans forward, blocking the light from Jesus—a strong visual metaphor.

Chinese painting seeks spiritual resemblance rather than physical likeness, which in terms of light and shadow means not simulating a single light source but expressing an overall conceptual sense of light through ink density variations. In light and shadow function, heavy ink expresses the orientation of mountain surfaces, while negative space expresses clouds and water light, creating a rhythmic sense of virtual-real interplay and light-dark alternation. In the light-dark system, Fan Kuan uses thick ink layered and accumulated to create rich light-dark

gradation, achieving “five colors from ink.” The sunny side of rocks uses lighter ink, while the shady side and ravines use extremely heavy burnt ink. This light-dark balance treatment makes the mountain body appear exceptionally solid and heavy with strong volumetric sense.

2.5 Labels for Color Relationship Generation in Artistic Images

Color is the most emotionally direct element in painting. PL-KG’s color system integrates physical attributes (hue, brightness, saturation), psychological perception, cultural symbolism, and artistic harmony principles, aiming to structure the complex color selection process.

Although the mural’s colors have severely faded due to damage, its color strategy can still be analyzed from restored versions and early copies. Da Vinci employed relatively subdued medium-purity tones and medium-brightness modes to suit the solemn atmosphere of the dinner scene. Jesus wears a red robe and blue mantle—red symbolizing sacrifice and divine love, blue symbolizing divinity and truth—color symbolism with specific meaning in Christian iconography. The disciples’ clothing colors vary, but through unified tones and adjacent color harmony, overall picture harmony is maintained.

Fan Kuan’s *Travelers among Mountains and Streams* embodies Chinese painting’s unique “five colors from ink,” as Zhang Yanyuan stated in *Records of Famous Paintings Through the Dynasties*: “Transport ink and the five colors are complete” [?]. From the heavy ink of foreground trees, to the dark ink of middle-ground mountains, to the light ink and negative space of distant mountains and clouds, a complete spectrum from low to high brightness mode is formed. This ultimate application of thick, thin, dry, wet, and black ink reaches directly to its inherent “spiritual resonance” and “truth,” embodying the Daoist philosophy that “the Dao gathers in emptiness” [?].

2.6 Labels for Brushwork and Texture Generation in Artistic Images

Brushwork and texture reveal the artist’s unique temperament and expressive techniques, serving as the most direct carrier of emotion and style [?]. PL-KG’s innovation lies in integrating Western painting’s emphasis on shaping and texture with Chinese painting’s emphasis on calligraphic quality and spiritual resonance into a unified analytical framework.

Da Vinci, to achieve oil painting-like delicate transitions and light-shadow effects on walls, abandoned traditional fresco techniques and adopted an experimental mixed tempera and oil medium applied directly to dry plaster. This exploration of material adaptation aimed to realize his famous “Sfumato” effect—blurring contours through extremely soft transitions to make figures seem to merge into light and air. This technique can be classified as a block brushwork pursuing ultimate soft edges.

Using Fan Kuan’s *Travelers among Mountains and Streams* as an example:

in terms of texture expression within volume shaping, Fan Kuan' s creation of "raindrop texture stroke" effectively shapes the volume and hard surface texture of northern mountain rocks. In conveying natural organic shapes, compared to mountains and forests, Fan Kuan' s lines depicting figures and animals are more flexible and relaxed, expressing the subjects' vitality [?]. In terms of positive and negative form, the majestic, solid, heavily inked mountain body constitutes the picture' s positive form, while the cleverly used negative space around it suggests flowing clouds and misty atmosphere, blending virtual and real harmony into the overall stable heaviness [?].

2.7 Labels for Edge Relationship Generation in Artistic Images

Edge treatment is an extremely subtle technique in painting that profoundly affects a picture' s sense of space, light, and even artistic style. PL-KG' s originality lies in treating edge relationships as an independent analytical dimension and extracting the fundamental differences between Chinese and Western edge treatment philosophies, thereby providing a unified analytical framework for these two different edge philosophies.

Western edges change with light and shadow. In the Western realist tradition, an edge' s clarity or blur 主要取决于光照条件. Bright areas, illuminated surfaces typically have clearer, harder edges (hard edges), while dark areas, shadowed surfaces have blurrier, softer edges (soft edges) to simulate light scattering' s physical effects. Masters like Rembrandt and Sargent are exemplars of using this light-shadow logic to shape form and space.

Chinese edges change with artistic conception. In Chinese painting, edges (i.e., lines) themselves are independent aesthetic objects, their treatment entirely serving the picture' s overall "spiritual resonance" and "artistic conception." Artists may use clear outlining lines to emphasize mountain structures and calligraphic quality, or use rendering and broken ink techniques to blur contours, creating moist, misty atmospheres. The logic is conceptual and poetic, not physical.

PL-KG structures these two philosophies and adds the sub-item "Chinese-Western fusion edge treatment," providing support for analyzing more complex edge languages in cross-cultural artworks.

3 PL-KG

This chapter details how the seven dimensions of PL-KG specifically connect to contemporary painting theory discourse. The discussion proceeds through composition generation, shape and form generation, perspective and space generation, light and shadow generation, color relationship generation, brushwork and texture generation, and edge relationship generation.

3.1 Composition Generation

As the “backbone of a painting”[?], composition affects both visual structure and carries artistic intent and emotion. PL-KG establishes “composition generation” as the first part, deconstructing it through multiple levels into a macro-to-micro system rooted in art theory yet possessing scientific feasibility and forward-looking computational implementation.

The primary function of composition is to carry and convey the work’s theme and emotion, with solid art theory support. Hu Wenxi et al. [?] note that the spatial characteristics presented by composition are the primary features of artistic image creation, linking compositional choices to artists’ emotional expression tendencies and providing basis for thematic delivery objectives. Roberts [?] emphasizes that composition is the “backbone of a painting,” giving pictures coherent rhythmic sense and solid structure through graphic element arrangement.

“Composition generation” includes seven secondary categories: “composition core objective,” “composition type,” “picture fullness,” “viewpoint structure,” “visual guidance,” “visual balance,” and “rhythm and cadence,” each further subdivided to fourth or fifth levels.

This study identifies four unique composition types: “full-frame composition,” “negative space composition,” “scattered-point composition,” and “focal composition,” with the first two classified by “picture fullness” standard and the latter two by “viewpoint structure” standard. “Visual balance” is subdivided into “color balance,” “light-dark balance,” “size balance,” “dynamic balance,” and “position balance.”

This study argues that visual balance is an objective phenomenon explainable through visual psychology and composition principles, not purely subjective feeling. Arnheim [?] points out that form shaping depends on the steepness of “light-dark gradient” gradients and emphasizes that “brush momentum” can build internal tension and strengthen volumetric sense. Johannes Itten [?] adds that warm-cool color contrasts can shape volume because they inherently imply spatial distance. Roberts [?] highly values “light-dark mass,” stating its role surpasses other picture areas, and that carefully arranged light-dark color blocks can transform a painting, while his mention of “asymmetrical layout” and visual attraction also provides examples for “dynamic balance” and “position balance.”

“Rhythm and cadence” is the temporal dimension of composition, affecting visual flow and emotional expression. Arnheim [?] emphasizes the dynamic nature of visual experience, discussing how shape repetition and variation constitute visual rhythm, supporting the “shape rhythm” classification. Hu Wenxi et al. [?] discuss color rhythm and brushwork’s effect on emotion and rhythm. Roberts [?] proposes “using rhythm to divide the picture,” affirming its status as a fundamental composition principle, and metaphorically states that good composition, like a song, has “rhythm and cadence,” confirming its validity as

a core dimension.

This study classifies “brushwork rhythm” under “rhythm and cadence,” alongside “color rhythm” and “shape rhythm.” Traditionally, detailed brushwork analysis is mainly used for artwork authentication or artist style research, but this study’s inclusion of it in the composition generation category provides a new entry point for art style and brushwork-driven artistic image annotation. Roberts’ [?] discussion of the relationship between brushwork and artist personality expression also supports this innovation—brushwork is not only a technical trace but also a direct carrier of emotion and rhythm.

3.2 Shape and Form Generation

Shape and form generation is the core of pictorial modeling, concerning not only the physical authenticity of objects but also carrying artists’ subjective expression and style construction. This knowledge graph constructs a complete theoretical framework from concrete representation to abstract expression around four dimensions: “volume shaping,” “shape type,” “shape treatment,” and “positive and negative form.”

Volume shaping aims to create visually three-dimensional images on a two-dimensional plane, focusing on precise expression of volume, structure, and texture. First, volumetric shaping is the foundation of volume shaping. Arnheim foundationally states: “The gradient of light and dark on a surface is key for the eye to perceive volume and three-dimensionality” [?]. This scientific principle reveals the essence of light-shadow modeling. Lang Yuqing [?] further validates this through art history cases: da Vinci first elevated light-shadow from intuitive practice to systematic theory, with chiaroscuro’s primary purpose being “scientific realism” and “three-dimensional volume”; Baroque painter Caravaggio pushed light-shadow’s modeling function to extremes through radical tenebrism, “using extreme contrast of light and dark” to emphasize object facets, making figures protrude like sculptures.

Second, light-shadow is a powerful tool for guiding gaze and constructing picture rhythm, also “an important tool for conveying emotion, creating atmosphere, and narrative in Western classical oil painting” [?]. Its narrative function’s core mechanism is gaze guidance. Caravaggio in *The Calling of Saint Matthew* precisely concentrates light on Christ’s finger pointing to Matthew, while surrounding figures fade into darkness, thereby “isolating the narrative moment and strengthening its emotional impact.” This perfectly illustrates the operational logic of “light area focusing gaze.” Rembrandt developed more complex rhythmic syntax, creating jumping light points through officers’ golden sashes, girls’ white collars, and musket reflections in *The Night Watch*. These “rhythms of alternating light and dark” guide gaze circuitously through the picture, giving static group portraits dynamic energy.

Finally, light-shadow is key to creating atmosphere and emotion. Da Vinci pioneered treating light-shadow as “key elements for creating picture atmosphere

and expressing emotion” in his notes, proposing three tonal prototypes: low-key painting creates “dramatic, profound, or mysterious atmosphere” through strong contrast; high-key painting conveys “bright, harmonious, serene effects” through large bright areas; middle-key pursues soft transitions for “maximum realism” [?]. Lang Yuqing’s cross-genre research empirically validates this theory: Caravaggio’s dark manner creates strong drama, Georges de La Tour’s candlelight scenes create mysterious gloom, while Chardin uses diffused, restrained light to achieve tranquil poetic expression [?]. Guo Maolai ultimately elevates this experience to “brightness mode” theory, providing modern terminology support for the four atmosphere categories in the knowledge graph [?].

3.3 Perspective and Space Generation

Constructing three-dimensional spatial illusion on a two-dimensional plane is a challenge in painting art. This study systematizes spatial generation methodology, integrating aerial perspective based on physical optics, spatial hierarchy processing based on compositional levels, Chinese and Western linear perspective systems based on geometry, and special spatial expression methods beyond convention, aiming to provide a comprehensive annotation framework for these complex techniques to support machine reading.

Aerial perspective, also known as atmospheric perspective, expresses spatial depth by simulating atmosphere’s effect on light and vision. As early as the 15th century, da Vinci systematically elaborated on this [?]. His notes explicitly record: “Through atmospheric differences, people can distinguish different distances of buildings,” noting that “more distant buildings should be painted with blurrier outlines and more blueness.” This incisive observation directly identifies two core visual changes caused by increasing distance: blurred outlines (clarity change) and blue color shift (warm-cool change), while color vividness also decreases (purity change). This provides solid historical and theoretical support for PL-KG’s decomposition of aerial perspective’s core elements into three dimensions: warm-cool change, clarity change, and purity change.

The scientific basis of aerial perspective lies in physical optics and visual psychology. Da Vinci’s observations already revealed its physical cause: “Air itself has physical properties that affect objects’ visual presentation with increasing distance” [?]. Modern science explains this as due to atmospheric particles’ light scattering (Rayleigh scattering makes distant views appear blue) and moisture’s occlusion of details. Psychologically, Arnheim’s [?] perceptual theory proposes that human vision has a tendency to simplify and organize—blurry, desaturated distant objects are naturally judged as being in deeper space.

In AI creation, applying aerial perspective principles is crucial. By learning color warm-cool gradients, outline clarity, and color purity gradations, we can effectively quantify how artists create three-dimensional spatial illusion with depth on two-dimensional planes, thereby enhancing AI’s ability to capture spatial sense and realism.

Linear perspective is the core technique for reproducing three-dimensional space on two-dimensional planes based on geometric principles [?]. This study creatively juxtaposes Western and Chinese perspective systems and proposes integrated application possibilities.

The Western perspective system was systematized during the Renaissance. Alberti defined it as a scientific representation method [?], while Piero della Francesca's "proportional diminution" mathematical principle provided an operational framework for perspective space calculation [?]. Da Vinci [?] not only studied standard one-point and two-point perspective but also keenly observed and recorded that when the field of view is extremely large, straight lines at the picture's edges appear curved—an early description of fisheye perspective effects. This system uses strict mathematical calculation to determine objects' size, position, and shape changes in pictures, providing powerful scientific tools for Western realist painting development.

In stark contrast is Chinese painting's unique spatial representation system, not based on fixed geometric principles but emphasizing expression of space and artistic conception perceived by the mind. Guo Xi in *Lofty Ambitions of Forests and Springs* proposed the "Three Distances method" (high distance, deep distance, level distance), systematically summarizing three basic spatial paradigms for Chinese landscape composition [?]. "Scattered perspective" allows painters to freely move viewpoints within the same picture, organically combining scenes from different times and spaces to create a "mobile viewpoint" visual experience, as in *Riverside Scene at Qingming Festival* [?]. Shi Xiaolan analogizes this perspective to film's lens language, profoundly explaining its unique narrative function [?].

Multiple dimensions clearly show the differences between the two systems. First, in theoretical foundation, Western linear perspective is strictly rooted in Euclidean geometry and optical principles, pursuing scientific precision [?]; Chinese conceptual perspective derives more from Daoist philosophy, painters' keen observation, and subjective memory, emphasizing comprehension of nature's spirit [?, ?, ?]. Second, this theoretical difference directly leads to different viewer positions and spatial goals. Western perspective sets a fixed, single viewing position forming a "visual cone" [?], aiming to create rational, measurable, illusionistic 3D space on 2D planes. Conversely, Chinese scattered perspective adopts mobile, multi-viewpoint viewing—"changing scenery with each step" [?]-aiming to evoke immersive, experiential, emotional space, namely "artistic conception." Third, in narrative function, Western linear perspective tends to capture a frozen moment like a "snapshot," fixing specific scenes. Chinese perspective, especially in scrolls, can unfold continuous narrative in time-space [?], like unrolling a "scroll" to lead viewers on a journey.

This study not only compares the two systems but also proposes the forward-looking category "Chinese-Western perspective fusion application." It 主张 that when analyzing modern creations, we can identify how artists combine Western linear perspective's "proportional accuracy" with Chinese perspective's "nar-

rative and artistic conception sense.” For example, using Western perspective for core subjects to ensure realism while employing scattered perspective for backgrounds to increase spatial flexibility and artistic conception. This fusion breaks through single perspective systems’ cultural limitations, bringing new possibilities to artistic image representation dimensions.

In modern art, perspective issues have moved toward more universal spatial problems [?]. Beyond conventional perspective systems, art history contains multiple special spatial expression forms based on specific cultural concepts or psychological experiences. This study categorizes them into “Chinese painting negative space,” “Western surreal space,” and “folk art planar space” for effective annotation of these non-traditional spaces.

Chinese painting’ s negative space significance extends beyond compositional technique. Wang Binggen’ s research shows Chinese painters use intentional blank space to stimulate viewers’ imagination, expressing the philosophical thought of “virtual and real mutual generation” [?]. Southern Song painters Ma Yuan and Xia Gui used large negative spaces in landscape paintings to create ethereal, profound artistic conception. This utilizes Arnheim’ s described visual “closure” tendency—viewers actively mentally supplement incomplete shapes or meanings, thereby participating in the work’ s completion [?].

Western surreal space 致力于探索潜意识和梦境世界. Artists like Dalí often use extremely realistic techniques to depict illogical, 超越现实的 scene combinations, such as the softened clocks in *The Persistence of Memory* [?]. This space construction relies on skilled application of traditional realistic techniques (like perspective, light-shadow) to depict seemingly real yet absurd scenes, creating a psychological dislocation between reality and dream that challenges conventional reality cognition.

Folk art’ s planar space is a more unique visual grammar based on philosophical functions. Jin Zhilin [?] explicitly states in his research that folk art’ s planar treatment is “direct interpretation of cosmic ontology’ s philosophical symbols.” Creators are not expressing personal feelings but translating cosmic laws and order. For example, Shaanxi paper-cut’ s “yin-yang fish” pattern, by completely eliminating light-shadow, perspective, and other realistic elements, transforms objects into pure symbols carrying the philosophical concept of “yin-yang harmony.” This planarization, by deleting accidental visual details, gives symbols high cognitive efficiency, conforming to visual communication’ s cognitive economy principle.

This study defines this planarization as a carrier of philosophical symbols, breaking through traditional cognition that treats it merely as a stylistic technique, providing a key annotation path for visual translation of cultural genes in datasets.

3.4 Light and Shadow Generation

Light and shadow are not only physical phenomena reproducing the objective world but also core languages in painting for shaping form, guiding narrative, and creating atmosphere. This study deconstructs light and shadow application into three levels— “function,” “light source,” and “light-dark system” —forming a complete logical chain from “why use light” to “where light comes from” to “how to light,” providing theoretical basis for refined annotation of light-shadow effects.

Light and shadow’ s functions in painting are multidimensional. This knowledge graph 归纳 s them into three core functions: “shaping volume and structure,” “guiding gaze and building rhythm,” and “creating atmosphere and emotion.”

First, “shaping volume and structure” is light-shadow’ s most basic function. Arnheim foundationally points out: “The gradient of light and dark on a surface is key for the eye to perceive volume and three-dimensionality” [?]. This scientific principle reveals the essence of light-shadow modeling. Lang Yuqing [?] further validates this through art history cases: da Vinci first elevated light-shadow from intuitive practice to systematic theory, with chiaroscuro’s primary purpose being “scientific realism” and “three-dimensional volume” ; Baroque painter Caravaggio pushed light-shadow’ s modeling function to extremes through radical tenebrism, “using extreme contrast of light and dark” to emphasize object facets, making figures protrude like sculptures.

Second, light-shadow is a powerful tool for guiding gaze and building picture rhythm, also “an important tool for conveying emotion, creating atmosphere, and narrative in Western classical oil painting” [?]. Its narrative function’ s core mechanism is gaze guidance. Caravaggio in *The Calling of Saint Matthew* precisely concentrates light on Christ’ s finger pointing to Matthew, while surrounding figures fade into darkness, thereby “isolating the narrative moment and strengthening its emotional impact.” This perfectly illustrates the operational logic of “light area focusing gaze.” Rembrandt developed more complex rhythmic syntax, creating jumping light points through officers’ golden sashes, girls’ white collars, and musket reflections in *The Night Watch*. These “rhythms of alternating light and dark” guide gaze circuitously through the picture, giving static group portraits dynamic energy.

Finally, light-shadow is key to creating atmosphere and emotion. Da Vinci pioneered treating light-shadow as “key elements for creating picture atmosphere and expressing emotion” in his notes, proposing three tonal prototypes: low-key painting creates “dramatic, profound, or mysterious atmosphere” through strong contrast; high-key painting conveys “bright, harmonious, serene effects” through large bright areas; middle-key pursues soft transitions for “maximum realism” [?]. Lang Yuqing’s cross-genre research empirically validates this theory: Caravaggio’ s dark manner creates strong drama, Georges de La Tour’ s candlelight scenes create mysterious gloom, while Chardin uses diffused, restrained light to achieve

tranquil poetic expression [?]. Guo Maolai ultimately elevates this experience to “brightness mode” theory, providing modern terminology support for the four atmosphere categories in the knowledge graph [?].

Systematic classification of light sources and rays is the prerequisite for controllable light-shadow annotation. This study divides light sources from three dimensions: direction, type, and quality/hardness.

Light source direction (e.g., side light, top light, backlight) determines objects’ light-dark distribution and volume presentation. Da Vinci’ s systematic exploration of light source direction was a revolutionary contribution to his light-shadow research [?]. Classic works in art history provide examples for different directional light functions: Caravaggio’ s side light in *The Conversion of Saint Paul* creates ultimate dramatic volume; da Vinci’ s top light in *The Last Supper* imparts sacredness; the backlight in *Ginevra de’ Benci* background strengthens spatial hierarchy. Thus, light source direction is closely linked to light-shadow function: side light excels at shaping volume, top light at creating atmosphere, backlight at building space.

Light source types are divided into natural light, artificial light, and subjective light. Guo Maolai distinguishes physically between direct light (e.g., sunlight) and transmitted light (e.g., candlelight) [?], corresponding to natural and artificial light. Lang Yuqing supplements the subjective light type beyond physical laws: the “unidentifiable warm yellow light” in Rembrandt’ s late self-portraits belongs to “emotional light” ; Caravaggio’ s direct strong light is “dramatic light” ; lighting violating physics in surreal paintings is “surreal light source” [?]. This classification’ s scientific basis is that natural and artificial light follow predictable optical laws, while subjective light serves artists’ emotional expression, such as Rembrandt using light to reveal “the inner state of depicted subjects.”

Light quality/hardness (soft/hard light) affects picture texture and atmosphere. Da Vinci earliest compared aesthetic differences between different light qualities: “harsh sunlight” can shape clear structure but loses details, while “overcast soft light” can present subtle transitions but weakens volume [?]. Lang Yuqing’ s empirical analysis reveals its application in artistic practice: Dutch still life painters use “concentrated light” (hard light) to sharply depict texture, while Chardin uses “diffused, restrained light” (soft light) to make objects glow with velvet-like warmth [?]. Guo Maolai points out that light hardness directly affects atmosphere expression: hard light (high contrast, narrow transition) enhances drama, while soft light (low contrast, wide transition) is more suitable for creating tranquil atmosphere [?].

3.5 Color Relationship Generation

Color is painting’ s most emotionally expressive and visually impactful element. This study’ s color relationship annotation system is a multidimensional framework integrating physical attributes, psychological perception, cultural symbolism, and artistic harmony principles, aiming to structure and theorize the com-

plex color selection process for systematic analysis.

Color's three basic attributes—hue, brightness, and purity—are core dimensions for analyzing and applying color relationships. Guo Maolai provides incisive definitions of these attributes' functions: brightness is color's skeleton, determining the picture's 素描关系, like Rembrandt's dramatic light-shadow structure; hue is color's identity, defining color's basic appearance, like Van Gogh's *Sunflowers* signature chrome yellow; purity concerns emotional intensity, like the strong unease from the blood-red sky in Munch's *The Scream* [?]. Da Vinci's notes already revealed interactions between color attributes, observing environmental colors' influence on objects' inherent colors [?], laying the foundation for color relationship theory.

These attributes' perceptual effects are widely validated in artistic practice and have scientific validity. Regarding warm-cool, Arnheim [?] confirms warm colors (e.g., red, orange) have advancing sense while cool colors (e.g., blue, green) have receding sense, providing psychological basis for creating spatial depth through warm-cool gradients. Regarding brightness, high-brightness tones generally appear light, like Monet's *Water Lilies* light purple mist, while low-brightness tones appear heavy, like Goya's *Black Paintings* gloomy atmosphere [?]. Regarding purity, high-purity colors can trigger strong visual tension, like Matisse's *Dance* vermillion, while low-purity colors can create quiet, restrained atmosphere, like Morandi's gray-toned still lifes [?].

This study systematically classifies color attributes, providing annotation work with a structured color logic analysis tool that transforms color selection analysis from intuitive perception to a reasoned “structural decision.”

Color symbolism is often closely connected to specific cultural contexts [?]. For example, in traditional Chinese culture, red symbolizes celebration and power, like the Forbidden City's vermillion walls; in the West, purple often represents nobility and divinity, like the Virgin's purple robe in Titian's paintings. However, cross-cultural common symbols also exist, such as gold representing sacredness in both Buddhist murals and Byzantine art. Clarifying these commonalities and differences is crucial to avoid color misinterpretation in cross-cultural art analysis.

This study systematically 梳理's color's cultural symbolic meanings, separating “East-West common symbols” (e.g., gold, white, black) from “characteristic symbols” (e.g., Eastern yellow and purple, Western green and blue), building a clear symbolic selection guide for artwork annotation under cross-cultural backgrounds.

Color harmony is the artistic method of organizing picture color relationships to achieve “order, unity, and harmony.” Zhang Chao and Zhu Xiaojun [?] systematically propose four main paths to color harmony: hue-dominated harmony (e.g., using complementary contrast for physiological visual balance); brightness-dominated harmony (e.g., using mid-brightness stability to unify pictures); purity-dominated harmony (e.g., reducing purity to ease color con-

flicts); and area contrast harmony (adjusting color area proportions to coordinate strong contrasting colors).

These harmony theories are highly scientific and operable. For example, complementary harmony can be precisely operated based on the 180° contrast relationship on Goethe's color wheel, with Van Gogh's red-green opposition in *The Night Café* being a classic case [?]. Color harmony theory provides effective solutions for analyzing seemingly conflicting color combinations. Additionally, Wang Feng's research points out material effects on color harmony [?]: natural materials' (e.g., wood, stone) inherent colors need contrast to enhance their beauty, while artificial materials can be more freely adjusted.

This study incorporates material carriers' physical properties (e.g., oil painting's heaviness vs. ink's transparency) into the harmony system, filling the "material harmony" gap in harmony systems to provide more detailed tags for subsequent dataset annotation.

3.6 Brushwork and Texture Generation

Brushwork and texture are painting language's most expressive and personalized elements, direct traces of artists' physical execution, carrying emotional flow and spiritual imprint. This study systematizes this complex domain, deeply analyzing from three levels—brushwork function, morphology, and texture/quality generation—to provide objective annotation basis for these highly personal characteristics.

Brushwork plays multiple roles in painting, being both technical means and emotional medium and style symbol. Shen Lan [?] systematically summarizes brushwork functions into three levels, providing solid theoretical foundation for this knowledge graph's classification.

First, brushwork is the basic tool for expressing the objective world and shaping form and space [?]. Artists construct objects' structure and dynamics through different brushwork forms and directions. For example, Van Gogh uses clear, flowing brushwork directions to convey celestial motion in *Starry Night*, while Freud uses broad, stacked brushwork to depict solid muscle structure. These brushworks serve "form" shaping, powerfully strengthening form's expressive meaning.

Second, brushwork is the direct carrier of presenting subjective emotion and picture rhythm [?]. Different brushworks can trigger completely different visual psychological responses. The jumping, light brushwork in Klimt's works conveys the painter's unrestrained inner world, while Schiele's tense, 滞涩 brushwork allows viewers to glimpse his inner depression and anxiety. Brushwork's smooth or heavy quality lets viewers intuitively experience spiritual agitation or emotional heaviness, thereby 烘托 ing the picture's overall atmosphere and artistic conception.

Finally, brushwork is the key element forming personal style symbols [?]. When

brushwork application surpasses pure depiction and lyric functions to form a stable, highly recognizable technical language, it becomes “the external manifestation of the painter’s mind and personality.” Van Gogh’s flowing short-line brushwork, Seurat’s rigorous pointillist brushwork, have all become unique symbols of their artistic styles. These symbolic brushworks not only reflect artists’ unique understanding of light and color but also highlight their decisive role as conscious artistic creation in pictorial representation.

Brushwork forms are ever-changing, their developmental trajectory reflecting profound transformations in painting concepts. This knowledge graph systematically divides brushwork morphology into four categories: calligraphic brushwork, block brushwork, pointillist brushwork, and linear brushwork.

Xiong Wei [?] traces the developmental history of oil painting brushwork in his research, providing clear context for this classification. From Van Eyck to most 18th-century classical painters, the 普遍采用 multi-layer transparent painting method, intentionally pursuing “neat, fine, smooth” effects without brush traces, which can be classified as “restrained brushwork.” Entering the 19th century, Delacroix began “expressing passion with bold brushwork,” opening the era of “expressive brushwork.” Impressionist painters “used short, broken small brushstrokes to depict objects,” like Monet’s trembling water surfaces, creating so-called “trembling brushwork.” Van Gogh developed continuous jumping thick short strokes and wave-like spiral strokes on this basis. Neo-Impressionist painters Seurat and Signac further developed brushwork into highly regular small color dots, forming unique “dot brushwork.”

Additionally, Xu Jizhuang [?] points out from a Chinese-Western comparative perspective that Western painting primarily uses mass modeling (corresponding to “block brushwork”), while Chinese painting uses line modeling, relying on “calligraphic brushwork” to produce rich variations in thickness, density, speed, and pressure, displaying unique sense of force, movement, and rhythm. This brushwork with calligraphic meaning is 归纳 d as “calligraphic brushwork” in this knowledge graph. This morphological classification system not only covers major brushwork types in Western oil painting history but also incorporates Chinese painting’s unique brushwork aesthetics, possessing cross-cultural inclusiveness.

Texture is the physical texture of picture surfaces, while quality is the tactile visual feeling created through texture and other means. This study divides texture generation into three levels: material synthesis, physical texture, and visual texture.

Physical texture is directly formed through paint’s physical properties and tool operations. Painters often use paint overlapping, stacking, scraping, and other techniques to create relief-like 凹凸感. Titian pioneered the technique of “thick application in bright areas, thin dyeing in dark areas” [?], greatly enhancing light effects and picture vividness. Rubens also enhanced picture tension by thickening bright areas and weakening dark areas, with this impasto method

becoming his unique means for shaping object quality and expressing light-color [?]. Besides painting, artists often create unique textures using mixed materials, such as mixing sand, plaster into paint, or directly collaging on canvas [?].

Visual texture simulates specific objects' surface qualities through physical texture, brushwork, color, and other means. Monet in his *Rouen Cathedral* series successfully expressed the church stone wall's solid, rough visual texture through thick physical texture created by multiple paint layers [?]. Rembrandt particularly favored expressing texture, with Roger Fry commenting that he "explored these things with a kind of innocent and passionate sensory pleasure" [?]. Rembrandt used high light-dark contrast to express metal texture of jewelry and armor, used meticulous small strokes to depict fluffy heaviness of cloaks, and used smooth brushwork to express silk's lightness [?]. Velázquez even cleverly used canvas texture itself to simulate hair texture. This transformation from physical to visual texture is key to painting art's creation of "tactile" illusion.

3.7 Edge Relationship Generation

Edge relationship treatment is an extremely subtle and crucial technique in painting, not only defining shapes and contours but also profoundly affecting picture spatial sense, light sense, atmosphere, and even fundamental attributes of artistic style. This study systematically investigates edge relationships as an independent analytical dimension, revealing their core functions in "creating atmosphere," "guiding focus," "defining shape," and "creating space," and innovatively proposing fundamental differences between Chinese and Western edge treatment philosophies, providing possibilities for data annotation of this subtle technique.

Edge clarity is not static but a continuous spectrum from sharp and clear to blurred disappearance. Yang Sitao [?] points out in his discussion that edge lines "or soft, or sharp, or clear, or blurred" treatment reflects painters' core skills and directly serves artistic expression. This study systematizes this spectrum into four categories: hard edge, broken edge, virtual edge, and soft edge.

Hard edge—clear, well-defined contours—is commonly used in classical realist painting to shape solid forms, like Lucian Freud's "hard-edge contour lines shaping female bodies with strong meaning of healthy female form" [?]. In contrast, soft and virtual edges create specific textures or atmospheres through blurred, gentle transitions. For example, Renoir "intentionally paints facial contours very blurry to express the soft, warm quality of women's skin." Impressionist painting pushes this treatment to extremes, eliminating clear light-shadow distinctions and contour lines to capture light's flowing, changing effects [?]. Monet's late *Water Lilies* series uses flexible brushwork to break edges between objects, making scenes 浑然一体 with rich yet transparent layers [?]. Broken edges lie between, suggesting contours through discontinuous, jumping brushstrokes, common in Impressionist and freehand styles. This clarity spectrum establishment provides precise vocabulary for quantifying and annotating different styles'

edge characteristics.

Edge functions extend far beyond shape definition. Arnheim long ago pointed out that lines (edges' concretization) can not only depict contours but even "reproduce brightness, space, and air" [?]. This study 归纳 s edge functions into four aspects: creating atmosphere and light sense, guiding visual focus, defining shape and contour, and creating spatial sense.

In creating atmosphere and light sense, painters can 虚化 edges to 烘托 backlight effects [?] or use overall soft edge treatment to create hazy fog atmospheres. In guiding visual focus, edge clarity contrast is an effective means—typically making subjects' edges clearer than secondary objects' (subject hard edge, secondary soft edge) powerfully attracts viewers' gaze to core areas. In creating spatial sense, edge treatment follows the basic rule of "solid front, virtual back." Da Vinci's [?] aerial perspective theory includes clarity gradient changes—near objects clear, distant objects blurred. In painting practice, making foreground objects' edges relatively solid and sharp while middle and background edges 虚化 and soft greatly enhances picture' s front-back spatial depth.

Western edges change with light and shadow. Since the Renaissance, Western painting has largely followed observation and reproduction of natural light-shadow. Whether an object' s edge is clear or blurry 主要取决于光照条件. Bright areas, illuminated surfaces typically have clearer, harder edges, while dark areas, shadowed surfaces have blurrier, softer edges to simulate light scattering' s physical effects. Masters like Rembrandt and Sargent generally follow the principle of bright areas solid, dark areas virtual, using dramatic light-shadow to shape distinct contours. Although Impressionist painters blurred contours, their purpose remained more truly reproducing light' s flowing, trembling visual effects on object surfaces [?, ?]. Thus, Western edge treatment' s fundamental principle is "changing with light-shadow," with perceptual, physical logic.

Chinese edges change with artistic conception. In Chinese painting, edge treatment surpasses objective light-shadow imitation, completely serving the picture' s overall "spiritual resonance" and "artistic conception." As Huang Binhong stated: "Ancient brushwork, today' s called lines, painting' s origin is based on this" [?]. Lines (edges) themselves are independent aesthetic objects. How to treat objects' edges depends on the artistic conception the artist wants to convey. For example, Wang Meng and Ni Zan use clear outlining to emphasize mountain structures and brushwork' s calligraphic quality; Ming painter Sheng Maoye uses ink wash rendering and loose brushwork to blur contours, creating moist atmosphere of rain impact [?]. Thus, Chinese edge treatment' s fundamental principle is "changing with artistic conception," with conceptual, poetic logic.

This knowledge graph' s core innovation in edge relationship theory lies in 提炼 and 概括 these two 截然不同的 edge treatment philosophies through comparison of Chinese and Western painting traditions, and creatively proposing the sub-

item “Chinese-Western fusion edge treatment.” This not only provides profound theoretical perspectives for understanding and analyzing artworks from different cultural backgrounds but also provides quantitative description frameworks for edge treatment modes in emerging artistic styles (like new Chinese oil painting, modern ink mixed media), holding important theoretical and application value.

Conclusion

The artistic image knowledge graph constructed in this study is essentially the concrete implementation of the “visual knowledge structurization” proposed in *On Visual Knowledge* [?] in the art domain, supplementing current art image knowledge graphs’ deficiencies in AI readability, usability, and scalability.

This study systematically demonstrates the theoretical rationality, scientific foundation, and core innovative value of PL-KG specially constructed for artistic image language and machine learning. Through layer-by-layer deconstruction of seven core dimensions—composition generation, shape and form generation, perspective and space generation, light and shadow generation, color relationship generation, brushwork and texture generation, and edge relationship generation—this study shows that this knowledge graph is not a simple pile-up of art terms but a rigorous, self-consistent, and forward-looking theoretical framework with “Chinese-Western fusion as cultural core, dataset annotation as practice orientation, and highlighting artistic image uniqueness as core goal,” providing key support for deep analysis and standardized annotation of artistic images.

The knowledge graph’ s core contributions manifest at three levels:

First, at the cultural theory level, it breaks through traditional art analysis limitations of “Western centrism” or “Eastern isolationism,” achieving systematic integration and theoretical reconstruction of Chinese and Western artistic languages. The graph is not a simple superposition of two systems but, starting from underlying logic, finds complementarity between Chinese and Western art at multiple dimensions like perspective, edge treatment, and brushwork, building a unified analytical framework: in perspective and space dimensions, it juxtaposes Western geometry-based linear perspective with Chinese philosophy-based “Three Distances method” and “scattered perspective,” not only clarifying their essential differences of “fixed viewpoint vs. changing scenery” and “rational space vs. emotional artistic conception,” but also innovatively proposing “fusion application” tags suitable for cross-cultural modern art creation; in edge treatment dimensions, it 提炼 s the core distinction between “Western edges changing with light-shadow” and “Chinese edges changing with artistic conception,” adding “Chinese-Western fusion edge treatment” sub-items to provide annotation basis for emerging styles like new Chinese oil painting; in brushwork and texture dimensions, it covers Western oil painting’ s block brushwork, pointillist brushwork, and Chinese painting’ s “texture method” and “ink method” calligraphic brushwork, while also attending to emerging cross-cultural styles like new Chinese oil painting and modern ink, filling cultural blind spots in existing

art analysis frameworks.

Second, at the practical application level, addressing the core pain points of large-scale art image dataset annotation—lack of unified standards, large subjective differences, and difficulty in quantification—this knowledge graph achieves the crucial transformation from qualitative description to structured annotation, providing actionable and reusable guidelines for annotation work. First, it solves definition ambiguity by transforming abstract qualitative concepts from traditional art into clear hierarchical tags. For example, it subdivides visual balance into five identifiable sub-items like color balance and dynamic balance, deconstructs rhythm and cadence into observable dimensions like color rhythm and brushwork rhythm, and even clarifies applicable scenarios for Chinese concepts like negative space and scattered perspective, giving annotators unified references. Second, it solves quantification difficulty by providing measurable reference indicators for core tags. For example, aerial perspective corresponds to three quantifiable dimensions of warm-cool change, clarity change, and purity change; visual guidance can be assisted by parameters like brightness gradients, perspective convergence point coordinates, and golden section point positions, transforming subjective aesthetic evaluation into objective, annotatable features. This structured design frees annotation work from dependence on personal experience, ensuring different annotators work within the same discourse system, fundamentally guaranteeing dataset consistency, accuracy, and depth, laying a high-quality data foundation for downstream tasks like artistic image creation intelligent generation.

Finally, at the artistic intelligence level, through targeted dimensional design, it precisely captures artistic images' core characteristics as a special visual medium. Unlike general images focusing on information recording, artistic images' core value lies in emotional expression, style transmission, and viewer interaction uniqueness. At the emotional expression level, the graph doesn't limit itself to physical attributes of light-shadow and color but emphasizes their function as emotional carriers. For example, in the light-shadow dimension, low-key painting corresponds to dramatic, mysterious atmosphere, high-key painting to serene, bright emotion; in the color dimension, distinguishing East-West common symbols and characteristic symbols enables annotation to cover the complete "color-emotion-culture" chain. At the style transmission level, the graph sets highly personalized elements like brushwork and texture as independent dimensions, emphasizing their uniqueness as "artist fingerprints," such as Van Gogh's flowing short-line brushwork, Seurat's pointillist brushwork, and Chinese painting's texture methods. These tags directly link to artists' style characteristics, while general image annotation rarely focuses on brushwork, making it difficult to reflect artistic creation's subjectivity. At the viewer interaction level, the graph focuses on artistic images' guidance and participation, such as movement line guidance in composition, Chinese negative space, and roaming experience of scattered perspective. These designs capture artistic images' characteristic of "actively guiding viewers' interpretation," while general images only pursue "clear information transmission" without interactive design intent.

Cultural Constitution emphasizes that “culture is a living organism, surviving in use and developing in innovation” [?]. Cultural elements not used in modern times are like “frozen living organisms” that need design innovation to revive them. This artistic image knowledge graph is both a systematic 梳理 of Chinese and Western art theories and an innovative attempt of “theory serving practice” : it breaks through cultural limitations through Chinese-Western fusion, solves annotation pain points through structured tags, and highlights artistic images’ essence through dimensional design, providing a crucial tool for constructing large-scale datasets with “deep artistic semantics.”

In the future, we can draw on *On Visual Knowledge*’s idea of “globally co-building visual knowledge dictionaries”[?], uniting art historians, computer graphics scholars, and AI researchers to further supplement visual tags for intangible cultural heritage arts, perfect the Chinese-Western fusion visual framework, and promote artistic visual knowledge from a dataset annotation tool to a foundational platform for AI art and cross-cultural research, providing key support for artistic image creation intelligent generation in the AI 2.0 era.

References

- [1] PAN Y. On Visual Knowledge (in English)[J]. *Frontiers of Information Technology & Electronic Engineering*, 2019, 8.
- [2] PAN Y. Cultural Constitution[M]. Beijing: Higher Education Press, 2011. ISBN: 978-7-04-025637-6.
- [3] Art History Teaching and Research Office, Central Academy of Fine Arts. A Brief History of Chinese Art[M]. Beijing: China Youth Press, 1990.
- [4] Foreign Art History Teaching and Research Office, Department of Art History, Central Academy of Fine Arts. A Brief History of Foreign Art (Revised Edition)[M]. Beijing: Higher Education Press, 2007.
- [5] Larousse Publishing Company, France, Michel Laclotte, Jean-Pierre Cuzin, DONG Q. Dictionary of Western Art[M]. Changchun: Jilin Fine Arts Publishing House, 2009.
- [6] HU W, WANG Z. Research on Oil Painting Art Style Classification Based on Multi-feature Fusion[J]. *Beauty and Times (Middle)*, 2023, (11):24-26. DOI:10.16129/j.cnki.mysdz.2023.11.015.
- [7] Sattarov Farxod. SCIENTIFIC-THEORETICAL FOUNDATIONS OF LEARNING THE RULES OF COMPOSITION AND DEPICTION IN PAINTING[J]. *Multidisciplinary Journal of Science and Technology*, 2024, 4(10):379-382.
- [8] Liu W. Exploration of the Minimalist Aesthetic in Geometric Form Language in Oil Painting[J]. *Highlights Design*, 2025, 9(1):11-14. DOI:10.54097/YZ7J8P34.
- [9] Kaidong W. The expression form and application research of order

- sense in modern oil painting creation[J]. *Frontiers Research*, 2025, 7(3). DOI:10.25236/FAR.2025.070307.
- [10] ARNHEIM R. Art and Visual Perception[M]. TENG S, trans. Chengdu: Sichuan People's Publishing House, 2019.
- [11] SHI X. Using Lens Language to Interpret Chinese Scroll Landscape Garden Painting: The Breakthrough of Scattered Composition on Two-Dimensional Static Continuous Space[J]. *Popular Literature and Art*, 2022, (20):32-34.
- [12] ROBERTS I. The Art of Composition[M]. SUN H, LIU H, trans. Shanghai: Shanghai People's Art Publishing House, 2012.
- [13] ITTEN J. The Art of Color[M]. YANG J, trans. Beijing: Beijing Science and Technology Press, 2021.07.
- [14] HE Z. On the Application of Artistic Anatomy in Painting[D]. China Academy of Art, 2010.
- [15] ZHENG L. Exploration of Structural Language in Modern and Contemporary Architectural Oil Painting[D]. Shaoxing University, 2024.
- [16] BAIDOLINO G, WANG Z. Methods for Appreciating Art Schools[M]. Beijing: Beijing Art and Photography Publishing House, 2016.06.
- [17] LIANG Z. Expression of Love and Death Themes in Viennese Secession Painting[D]. Guangzhou Academy of Fine Arts, 2023. DOI:10.27626/d.cnki.gzmsc.2023.000381.
- [18] Wolf N, Grosenick U. *Expressionism*[M]. Taschen, 2004.
- [19] ZHI Y. Induction and Simplification[D]. China University of Mining and Technology, 2023. DOI:10.27623/d.cnki.gzkyu.2023.001231.
- [20] Fry R. *Vision and design*[M]. Courier Corporation, 2012.
- [21] TAN J. Application of Picasso-Cubist Painting Language in Oil Painting Creation 'Image and Form' [D]. Harbin Normal University, 2024. DOI:10.27064/d.cnki.ghsu.2024.000373.
- [22] CHEN Q. Realist Discourse and Transforming Picasso: Translation and Introduction of Cubism in Early 20th Century[J]. *Art*, 2023, (01):91-99. DOI:10.13864/j.cnki.cn11-1311/j.007029.
- [23] ZHAI X. Research on Geometric Structural Modeling Language in Oil Painting Creation[D]. Guizhou Minzu University, 2024. DOI:10.27807/d.cnki.cgzmz.2024.000451.
- [24] DAI X. The Concept of 'Picturesque' and British Landscape Painting Art Style[J]. *Art*, 2009, (11):118-123. DOI:10.13864/j.cnki.cn11-1311/j.2009.11.016.
- [25] Kenneth C. *The Nude: A Study in Ideal Form*[M]. Princeton University Press, 1956.
- [26] ALBERTI L B. *Della pittura*[M]. Milano: Società Tipografica de' Classici Italiani, 1804.

- [27] DA VINCI L. *Leonardo da Vinci's Note-books*[M]. London: Duckworth and Company, 1906.
- [28] DELLA FRANCESCA P. *De prospectiva pingendi*[M]. Firenze: Sansoni, 1942.
- [29] HONG H. Alternative Interpretation of the Three Distances Method in Landscape Painting[J]. *Art Research*, 1998(4):13-19.
- [30] CAO S. Chinese Painting Scattered Perspective and Cézanne's Multiple Perspective[J]. *Art Work*, 2017(1):18-19.
- [31] COQUIOT G. *Paul Cézanne*[M]. Paris: DigiCat, 2022.
- [32] GREENBERG C. Modernist painting[M]//*Modern Art and Modernism*. London: Routledge, 2018:5-10.
- [33] FINKELSTEIN H. Dali's paranoia-criticism or exercise of freedom[J]. *Twentieth Century Literature*, 1975, 21(1):59-71.
- [34] WANG B. On the Meaningful 'Negative Space' in Oil Painting Works[J]. *New Art*, 2010, 31(3):51-53.
- [35] JIN Z. On Chinese Folk Art[J]. *Art Research*, 2003(3):63-65.
- [36] ZHU F. The Narrative of Zhang Zeduan's 'Riverside Scene at Qingming Festival' Painting Scroll[J]. *Literary and Art Criticism*, 2014, (04):114-116. DOI:10.16566/j.cnki.
- [37] LANG Y. Preliminary Exploration of Light and Shadow in Western Classical Oil Painting[J]. *Art Observation*, 2025, (06):129-131.
- [38] LI Z. 'Three Major Planes', 'Five Tones' and 'Light-Dark Boundary Line' [J]. *Art*, 1956, (09):62. DOI:10.13864/j.cnki.cn11-1311/j.1956.09.016.
- [39] GUO M. Introduction to Visual Art[M]. Beijing: People's Fine Arts Publishing House, 2000.
- [40] ZHANG J. Application of Color in Brand Image Design[J]. *Art Panorama*, 2018, (03):106-107.
- [41] ZHANG C, ZHU X. Application of Color Contrast and Harmony in Advertising Design[J]. *Journal of Northwest Normal University (Social Sciences Edition)*, 2011, 48(03):129-132. DOI:10.16783/j.cnki.nwnus.2011.03.021.
- [42] WANG F. Visual Embodiment of Design Material Aesthetics[J]. *Journal of Nanjing Arts Institute (Fine Arts and Design Edition)*, 2006, (04):199-200.
- [43] YANG S. On Edge Lines[J]. *New Art*, 1997, (03):31-32.
- [44] XIONG W. Development and Aesthetic Characteristics of Oil Painting Brushwork[J]. *Art Research*, 2005, (02):92-97.
- [45] LI F. Application of Backlight in Realistic Oil Painting[D]. Nanjing University of Aeronautics and Astronautics, 2009.

- [46] HUANG B. Huang Binhong on Art[M]. WANG Z, ed. Shanghai: Shanghai Calligraphy and Painting Publishing House, 2012:204.
- [47] CAHILL J. The Lyric Journey: Poetic Painting in China and Japan[M]. HONG Z, GAO S, GAO X, trans. Beijing: SDX Joint Publishing Company, 2012.
- [48] SHEN L. On the Spiritual Connotation of Oil Painting Brushwork[J]. Art, 2016, (06):106-108.
- [49] XU J. 'Linear Expression' in Chinese Painting[J]. Art Research, 1994, (02):39-43.
- [50] SUN Y. Research on Brushwork and Texture of Oil Painting[D]. China Academy of Art, 2012.
- [51] FRY R. Selected Art Criticism of Fry[M]. SHEN Y, trans. Jiangsu Fine Arts Publishing House, 2010.07.
- [52] ZENG J. On the Use of Light in Rembrandt's Oil Paintings[J]. Art, 1988, (10):18-19.
- [53] SHEN D. Research and Reference on Texture in Henk Helmant's Still Life Oil Paintings[D]. Fujian Normal University, 2023.
- [54] CHEN J. Research on Hyperrealistic Language in Still Life Sketching[D]. China University of Geosciences, 2013.
- [55] JIANG Y. Re-understanding Cubism[J]. Art, 2011, (06):118-123.
- [56] FENG Z. Cubism, Futurism, Abstractism[J]. General, 1928, 6(3):375-398.
- [57] JIANG Y. On Visual Perception and Spatial Construction in Painting[J]. New Art, 2007, (06):74-80.
- [58] Zhai W, Cao Y, Zhang J, et al. Exploring figure-ground assignment mechanism in perceptual organization[J]. *Advances in Neural Information Processing Systems*, 2022, 35:17030-17042.
- [59] WEI B. From Contained Posture to Simulated Image: The Rise of Modern Landscape and Ink Techniques[J]. Art Observation, 2024, (09):66-73.
- [60] LI C. History of Modern Chinese Oil Painting[M]. Shanghai: Shanghai Calligraphy and Painting Publishing House, 2007.12.
- [61] CHEN C. The Modernist Path of Lin Fengmian, Zhang Ding and Wu Guanzhong's Ink Painting Creation[J]. Art, 2011, (02):96-101.
- [62] WU H. Wu Hong's Collection: Chinese Painting from Five Dynasties to Southern Song[M]. Shanghai: Shanghai People's Publishing House, 2023.06.
- [63] ZHANG G, ZHAO D. Leonardo da Vinci and His Era[J]. World Art, 2016, (04):2-8.

- [64] TANG Y. A Brief Discussion on the Influence of Lao-Zhuang Aesthetics on the Creation of Song Dynasty Landscape Paintings: Taking Fan Kuan' s *Travelers among Mountains and Streams* as an Example[J]. *Open Access Library Journal*, 2024, 11:e12089.
- [65] JIANG Runyu. Reason and Nature—The Landscape Painting Embodiment of Taoist Morality[C]//Proceedings of the 2021 2nd International Conference on Mental Health and Humanities Education (ICMHHE 2021). Atlantis Press, 2021:252-255.
- [66] XU J. Art Dictionary[M]. Shanghai: Shanghai Lexicographical Publishing House, 2022.
- [67] YANG Y. Formal and Iconographic Analysis of ‘Travelers among Mountains and Streams’ [J]. *Art*, 2018, (01):134-135.
- [68] DU Z. Formal Order in Leonardo da Vinci' s Paintings: Taking ‘The Last Supper’ as an Example[J]. *Beauty and Times (Middle)*, 2025, (09):89-91.
- [69] ZHANG S. Expression of Dramatic Conflict Created by Light and Shadow in Oil Painting: Taking 17th Century European Oil Works as Examples[J]. *Beauty and Times (Middle)*, 2025, (08):96-98.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.