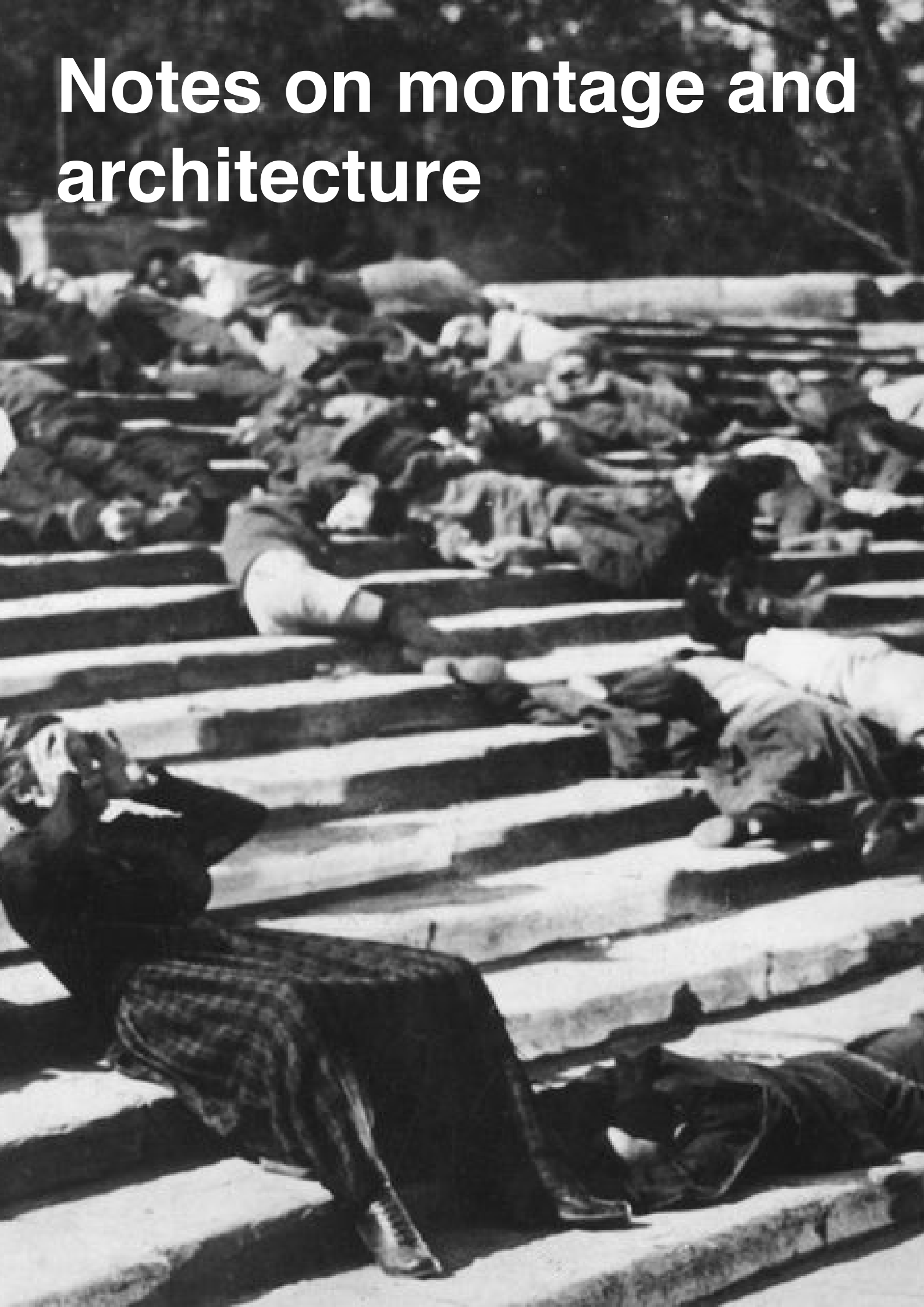
Notes on montage and architecture



Premise

As my engagement with architectural theory deepened, I increasingly found myself constrained by the limits of language. English, not being my first language, posed both a linguistic and conceptual barrier. My vocabulary felt inadequate, and my interpretive faculties often faltered when confronted with dense theoretical texts. In an attempt to navigate this gap, I began developing a parallel practice of making verbal and graphical notes—visual mappings, marginalia, and improvised lectures delivered to no audience but myself. These experiments, born of necessity, gradually revealed themselves as productive methods of inquiry. They transformed passive reading into a more active, embodied mode of interpretation.

In the course of this process, I also became attuned to a broader condition—one not limited to my own experience. Within architectural education, particularly in India, critical texts are often treated as secondary materials, referred to rather than engaged with, cited rather than interpreted. The absence of interpretive rigor—of a willingness to ask what a text does, rather than merely what it says—risks reducing architectural theory to footnotes or ornamentation. And yet, it is precisely through the act of interpretation that architecture might think itself otherwise.

This essay represents my first attempt to formalize that interpretive process. It engages with Montage and Architecture, a collaborative essay by Sergei Eisenstein, Yve-Alain Bois, and Michael Glenny, originally published in Assemblage (1989, MIT Press). The text is treated here not as a stable source of meaning, but as a provocation—an invitation to think across media, to trace the cinematic logics that animate spatial experience. What follows is not a scholarly exegesis, but a personal negotiation with ideas that have, in various ways, unsettled and reoriented my understanding of both cinema and architecture.

Nihal Galathia

Notes on Yve-Alain Bois forward

The following text and images are both verbal and visual notes compiled during a reading of Sergei Eisenstein's essay Montage and Architecture. The aim is not merely to reference the essay but to actively engage with it—to infer, interpret, and internalize its ideas as a mode of encountering architectural theory.

What follows may often resemble a simplified reconstruction of the original text, interwoven with personal interpretations. For me, the foreword by Yve-Alain Bois holds particular significance; it serves as a powerful pre-text, framing Eisenstein's writing in ways that render it both more accessible and more profound. The emphasis, therefore, rests heavily on Bois's foreword—not simply to unpack it, but to reflect upon it through the lens of one's own understanding.

Expanding on Bois's writing opens a field of interpretation that is both rich and generous. Even my own tentative or imprecise readings have proven productive, offering entry points into the kind of thought process required to grasp Eisenstein's ideas—and to reinterpret them in light of one's own experiences and inquiries.

Fittingly, the essay itself exemplifies the very theory it sets out to explore. With Bois's foreword acting as a montage-like prologue, the fragments of Eisenstein's argument acquire new coherence, becoming more legible and more resonant.

By way of background: Sergei Eisenstein, the influential Soviet filmmaker and theorist, wrote Montage and Architecture, which was published posthumously in Assemblage, Volume 10 (1989), by MIT Press.

It is perhaps best to introduce Eisenstein in his own words:

"It seems that all the arts, throughout the centuries, tended toward cinema. Conversely, cinema helps us to understand their methods."

This formulation—of a continuous critical loop, as opposed to a linear genealogy of "precursors"—offers a striking perspective. For Eisenstein (or at least in Bois's reading of him), earlier art forms do not merely serve as stepping stones toward cinema; rather, cinema re-illuminates them, opening new ways of seeing and understanding their structures. The relationship, in other words, is reciprocal—not unidirectional.
Eisenstein's own definition of montage deserves to be quoted at length:

"In themselves, the pictures, the phases, the elements of the whole are innocent and indecipherable. The blow is struck only when the elements are juxtaposed into sequential image."

This deceptively simple insight had a powerful historical arc. Following the Russian Revolution, Lenin recognized cinema's enormous potential as a medium for propaganda, agitation, and mass education. While traditional art—presented in opera houses and museums—remained largely the preserve of the elite, cinema emerged as the ideal artistic form for the world's first proletarian state.

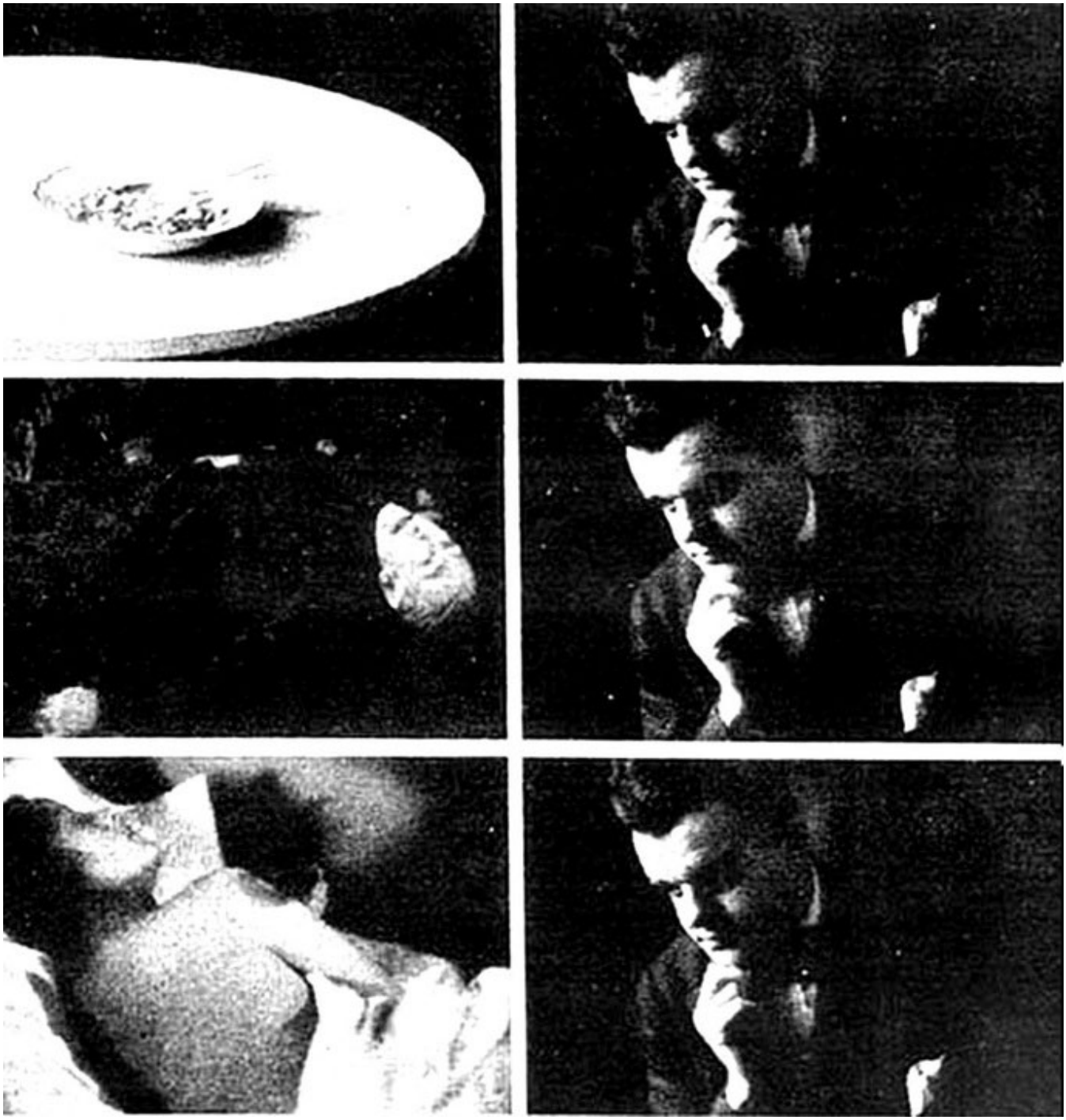
"Of all the arts, the most important for us is the cinema," -Vladimir Lenin

Lenin's pronouncement, and his efforts to nationalize the Soviet film industry, testify to an acute awareness of the medium's ideological and pedagogical power. Yet the Soviet Union faced material constraints: film reels were scarce. Faced with this shortage, students at the newly founded film institutes—including Eisenstein, Lev Kuleshov, and Vsevolod Pudovkin—

turned to analysis. They dissected films such as The Birth of a Nation and Intolerance by D.W. Griffith, studying them shot by shot, frame by frame. By rearranging these sequences, they began to explore how meaning could shift through the manipulation of order—thus giving birth to the theory of montage.

The concept became most vividly expressed in what is now known as the Kuleshov Effect: an expressionless face, when juxtaposed with different images—a bowl of soup, a child's coffin, a reclining woman—elicits entirely different emotional responses: hunger, grief, desire. A single image is inert; only through its relation to another does meaning emerge.

Though elementary, this insight is foundational. In its simplicity lies its radical power—one that is easily overlooked or forgotten in contemporary image-saturated culture.



Effekt Kuleshova, Semyon Raytburt, 1969

The idea of montage was not a fixed doctrine for Eisenstein, but an evolving concept—one he revisited and reinterpreted across his writings. In his essay Beyond the Shot, Eisenstein critiques the absence of dynamic cinematographic technique in early Japanese cinema. And yet, he simultaneously identifies within Japanese culture a deep-rooted affinity with the principles of montage.

Whether in the pictographic structure of kanji, the suggestive compression of haiku poetry, the layered compositions of woodblock prints, or the choreographic rhythms of Kabuki theatre, Eisenstein saw a shared aesthetic logic: meaning unfolding through juxtaposition, interval, and sequence. These traditions did not merely resemble montage—they embodied it.

Tite Kubo, Matsuo Basho, Utagawa Kuniyoshi, Toshusai Sharaku

Thus, even if Japanese cinema lacked formal cinematographic innovation at the time, Japanese culture, in Eisenstein's view, had long been practicing montage. The technique did not originate with film—it was already inscribed in the cultural fabric.

In confronting the critique levied by Kazimir Malevich—who accused Eisenstein's films of indulging in pictorialism, that is, privileging graphic composition over narrative continuity—it is useful to turn to a specific visual comparison. Malevich's painting The Knife Grinder bears a striking formal resemblance to one of Eisenstein's most iconic shots from Strike. The dynamic composition, fragmentation of form, and angular tension in both works appear to support Malevich's claim: Eisenstein's films, it seems, lean more toward the visual grammar of painting than the narrative logic of cinema.

However, Eisenstein's response inverts the accusation and, in doing so, opens up a more critical discourse. If his films resembled paintings, it was not because they were lacking in cinematic rigor, but because painters—consciously or not—had already been employing cinematic techniques long before the medium itself existed. In his view, montage was not exclusive to film; it was a broader cultural logic latent in many artistic traditions.

Strike (1925), Sergei Eisenstein & The Knife grinder (1912), Kazimir Malevich

Eisenstein repeatedly returned to this idea throughout his writings, identifying cinematic qualities in works that predated cinema by centuries and emerged from disparate cultural contexts. For instance, he found a proto-cinematic sensibility in El Greco's dramatic lighting and elongated figures—formal decisions that seemed to suggest movement and psychological intensity.

Opening of the fifth seal (1614), El Greco

Saint Jerome (1610), El Greco

He drew attention to the sequential logic in the compositions of Utamaro and other Japanese woodblock printmakers, whose visual storytelling unfolded across frames with deliberate rhythm.

Indeed, if one were to apply the same analytical method used by Soviet filmmakers—reordering frames to test how meaning shifts—to certain woodblock compositions, the principle of montage becomes vividly apparent. In one such example, a grid of six images arranged in three columns and two rows leaves the narrative sequence open-ended. The viewer's eye may move horizontally, vertically, or even diagonally, assembling the story in various possible ways. This multiplicity allows for interpretive freedom.

Men's Stamping Dance, Kitagawa Utamaro

In contrast, when the same images are reorganized into a graphical composition as below, the composition dictates a specific sequence, guiding the viewer along a singular interpretive path. The structure becomes didactic—a visual directive rather than an invitation.

A similar logic governs the murals of Mexican painter Diego Rivera, particularly in large-scale works like Dream of a Sunday Afternoon in Alameda Central. The sheer size of the mural prevents it from being taken in all at once; it must be read in fragments, across time and space. Moreover, its compositional structure—anchored by the Alameda Central park—deliberately stages Mexican history as a sequential, cinematic experience. The viewer moves across the surface as one might traverse frames in a film reel.

Dreams of Sunday afternoon in Alameda Central Park, Diego Rivera, 1947

Eisenstein also turned to the long, scroll-based paintings of Chinese and Japanese traditions. These works unfold gradually, with each unrolling section revealing a new narrative moment. Like cinema, they rely on time and movement to construct meaning. For Eisenstein, this temporal unfolding was another instance of montage embedded in a non-Western art form.

Japanese 19th Century Scroll Painting

Other examples further broadened his interpretive reach. Daumier's satirical compositions, for instance, captured bodily gestures and facial expressions that seemed to encode motion within stillness. Robert Delaunay's vibrant juxtapositions of color and geometry fractured and reassembled forms in ways that recalled cinematic editing. Soviet painter Valentin Serov used compositional strategies to suggest movement—motion suspended within the painted frame.

Counsel for the defense (1862-1865) and A scene from comedy (1858-1862), Honoré Daumier

Nu à la coiffeuse (1915) Robert Delaunay and A girl with peaches (1887), Valentin Serov

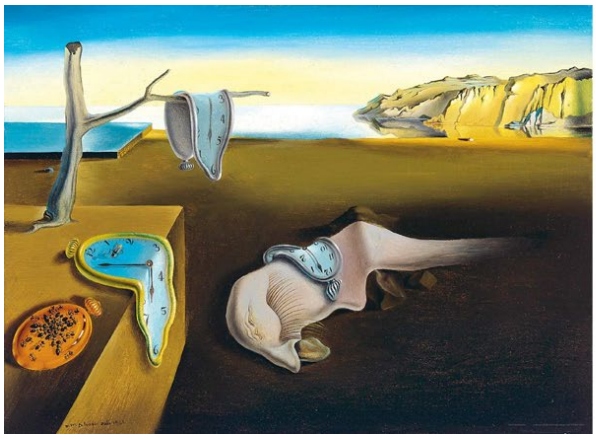
With each essay, Eisenstein's catalogue of "cinematographic" artists grew. As Yve-Alain Bois argues, this was neither a defensive maneuver to deflect Malevich's critique, nor a mere exercise in synesthesia—a search for correspondences between the arts. Rather, Eisenstein's observations revealed a deeper, more structural claim: that the principles of montage—sequentiality, rhythm, affective assembly—exist not only in cinema, but are embedded in the very grammar of visual art itself.

Eisenstein repeatedly sought a middle ground between the pictorial and the cinematic—neither painting nor film in isolation, but something in-between. He coined the term “cinematism” to describe this hybrid logic. Yet, as he himself suggests, this term is two-faced: if so many static visual artworks already exhibit dynamic, time-bound characteristics, are they still simply “paintings”? Or are they, in effect, already cinematic?

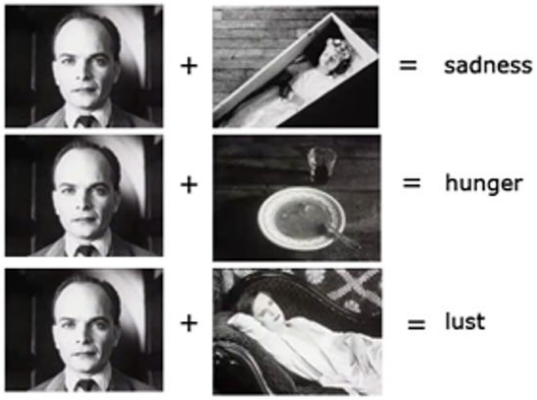
For Eisenstein, even a still image is structured through time—and perceived through time. This is the crux of cinematism: a concept with two interrelated aspects:

-Time inscribed within static form
  (Still images can suggest or embody movement or temporal unfolding)

-The sequential nature of aesthetic perception
  (Perception is not instantaneous but accumulative; context alters meaning)

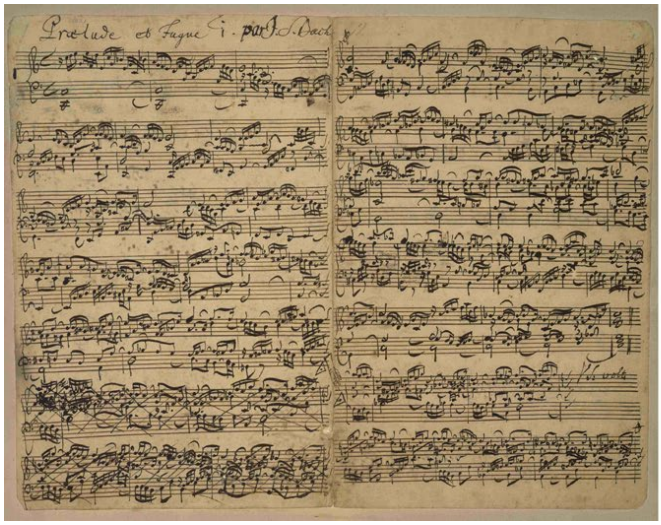


The Persistence of Memory (1931), Salvador Dali



Effekt Kuleshova, Semyon Raytburt, 1969

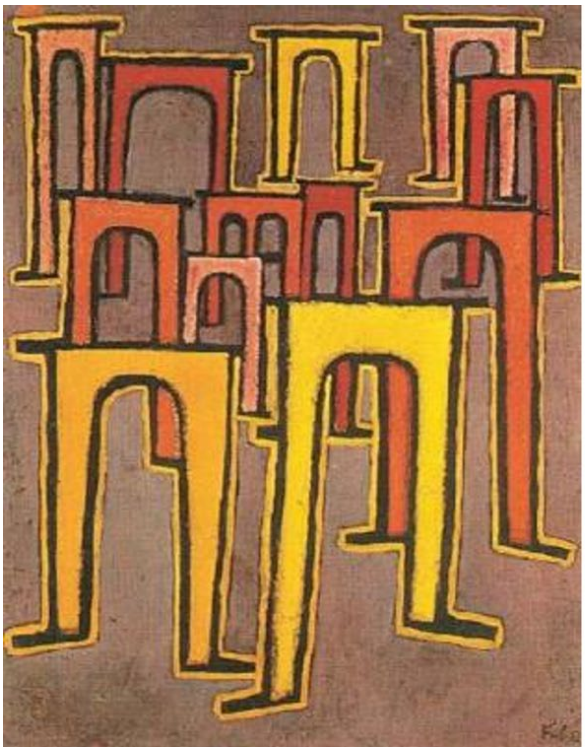
This position fundamentally challenges Gotthold Ephraim Lessing’s famous distinction between time-based and space-based arts. In his Laocoön, Lessing argued that music and poetry unfold in time, while painting, sculpture, and architecture are spatial—meant to be apprehended all at once. According to Lessing, time is essential for the former, while space is definitive of the latter.

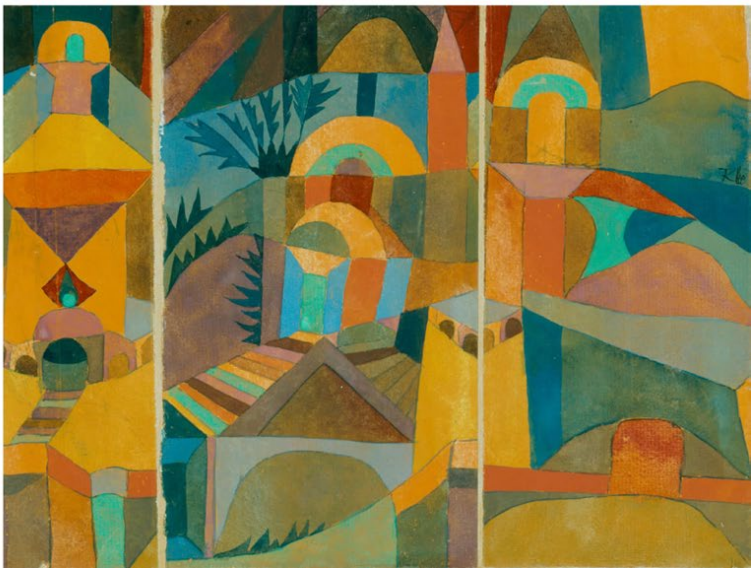


Prelude in C major, Bach, The Starry Night, Vincent Van Gogh & David, Michelangelo

Eisenstein’s understanding, however, undoes this binary. He contends that painting and sculpture also unfold over time, not necessarily in their making, but in the mode of perception. We do not grasp a painting in a single glance; rather, we “read” it—allowing our gaze to travel across its surface, assembling a composition progressively. Similarly, we encounter sculpture not from a fixed point but as a sequence of changing perspectives. These are temporal experiences embedded within spatial forms.

This insight also revises Immanuel Kant's claim that the experience of the beautiful is immediate—something that occurs in an instant, detached from temporal duration. Eisenstein's theory, by contrast, locates meaning in the temporal unfolding of perception. Beauty, in this sense, is not immediate, but constructed, as the viewer moves through or across a work. This argument echoes Paul Klee's idea of the eye "grazing" over the surface of a painting—isolating forms and rhythms, which are then assembled by the mind into a coherent whole. Klee foregrounds the active role of perception: the image is not simply seen, it is read.

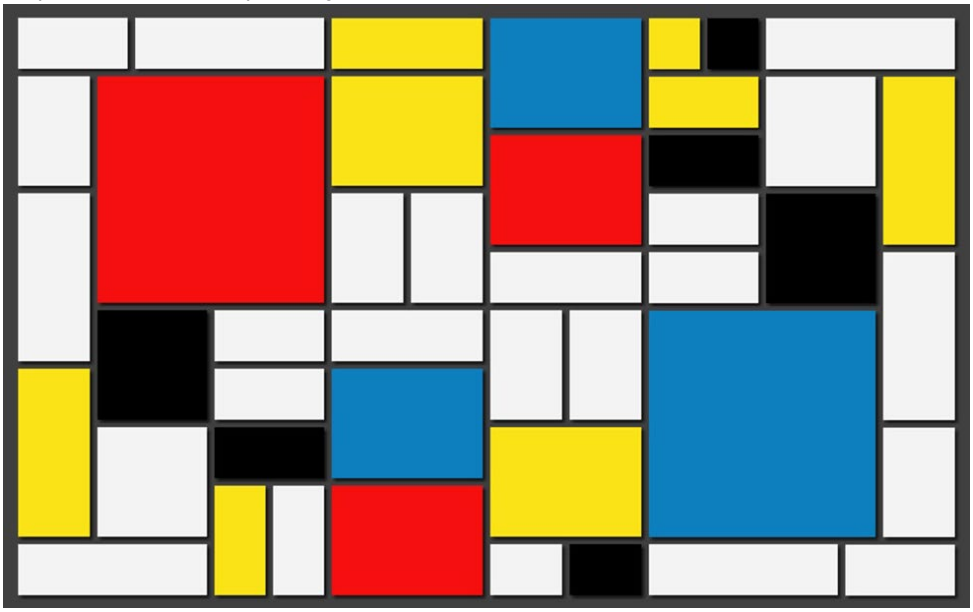




Revolution des Viadukts (1937) and Temple Garden (1920), Paul Klee

Even in virtual or two-dimensional space, the eye follows imaginary paths, guided by contrast, weight, directionality—paths that construct a kind of narrative, a sequence. These perceptual sequences are not random: they are shaped by the artwork's internal logic.

In contrast, one might consider Mondrian's compositions, which appear to privilege flatness and instantaneity. Yet even here, the act of looking is not purely simultaneous. Despite their graphic clarity, Mondrian's grids still solicit a kind of slow unfolding, as the eye moves from dominant forms to subtler variations. The painting, though ostensibly static, is still experienced temporally.



Composition with red, blue & yellow (1930), Piet Mondrian

When shifting from the virtual space of the image to the physical space of the built environment, Eisenstein continued to analyze through the lens of montage and movement. Yet, a critical distinction emerges: whereas in painting the eye travels across a surface, in architecture it is the body that moves through space. Perception becomes embodied. Architecture, unlike painting or cinema, is experienced through the choreography of the walking observer.

People matching artworks, Stefan Draschan

IIM Bangalore, B V Doshi

Yet for much of architectural history—and architectural education—this experiential dimension has remained largely overlooked. Architects are trained to conceptualize space through tools such as plans, sections, and elevations, representations that are fundamentally static and abstract. These are not how we perceive architecture in real life. We do not experience buildings in orthographic projection—we inhabit them temporally, through movement, light, shadow, and changing perspective.

This long-repressed phenomenological dimension of architecture resurfaced in the early 20th century, when Eisenstein's cinematic theories and Le Corbusier's architectural ideas began to converge around questions of sequencing, narrative, and spatial experience. For Eisenstein, this critical loop between cinema and architecture was even more profound than with painting. Unlike painters, Eisenstein was not only the son of an architect, but also trained as an engineer. His familiarity with architectural thinking gave him an intimate point of entry. Cinema helped him see architecture differently; and in turn, architecture reshaped his cinematic imagination.

In his films, Eisenstein persistently challenged the convention of treating architecture as a mere background—a passive setting for action. He sought instead to make the built environment an active agent of storytelling. Architecture was not to be filmed, but cinematically activated—imbued with rhythm, tension, and montage.

October (1927) and Ivan the terrible (1944), Sergei Eisenstein

In doing so, he developed what we might call a “cinematic judgment” of architecture. As he put it:

> *“Some buildings are more apt than others.”*

Yet, unlike cinema, architecture is non-mimetic. It does not represent reality; it is reality. And still, Eisenstein saw in it the same montage-like features that structured film: sequentiality and parallax—two essential components of spatial perception. Architecture unfolds as a sequence of visual and sensory events, and it changes in response to the movement of the observer. The built world does not remain still; it reorients itself in time, as our viewpoint shifts. We are not outside architecture, observing it as an image—we are within it, co-constructing its meaning through movement.

In this way, Eisenstein revived a too-often neglected “material” of architecture: the experience of movement itself. While architecture may be physically static, it is made dynamic through the perception of those who inhabit it.

A key reference here is Eisenstein’s renowned essay on Giovanni Battista Piranesi, whose Carceri etchings captured what Eisenstein termed a “vertiginous fragmentation” of space. These labyrinthine visions do not offer a fixed or unified view but generate a multiplicity of perspectives, dislocating the observer and inviting a fundamentally dynamic reading of space.

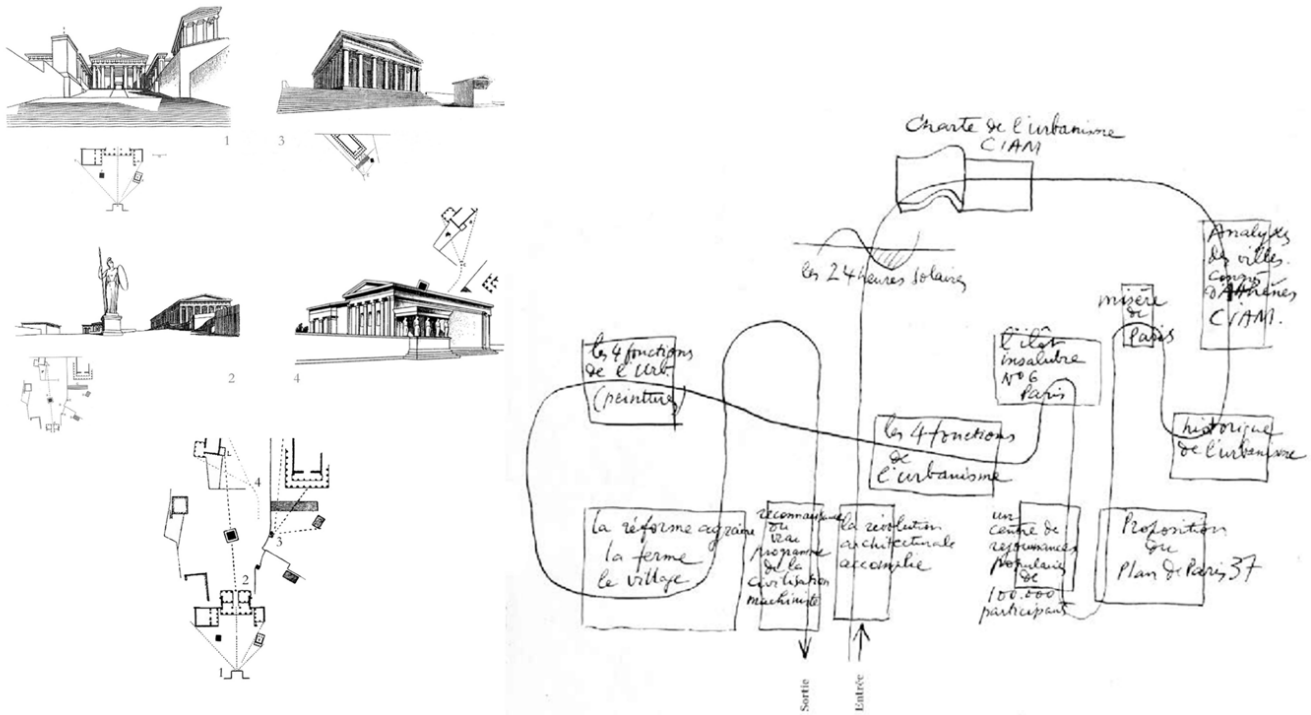
14th plate in the second edition of Le Carceri d’Invenzione (1761), Giovanni Battista Piranesi

Piranesi's compositions confound traditional architectural representation. They conflate plan, section, and elevation; they disrupt the singular point of view and instead invoke a spatial experience that is ambiguous, layered, and psychologically charged. In this, Piranesi anticipated what Eisenstein would later theorize through montage: a space not as a resolved whole but as an interpretative field—a shifting terrain of fragments.

Eisenstein's montage theory found a kindred sensibility in Piranesi's fragmented vision. And while Piranesi influenced 18th-century figures like Étienne-Louis Boullée and the English picturesque movement, the experiential dimension of architecture he proposed would recede from architectural discourse for nearly a century.

It was only in the late 19th century that Auguste Choisy reanimated this idea. In his Histoire de l'Architecture, Choisy offered a spatial analysis of Greek temples based on the sequence of visual experiences that unfold as one moves through space—not just the temple as an object, but as a narrative progression of thresholds, axes, and alignments.

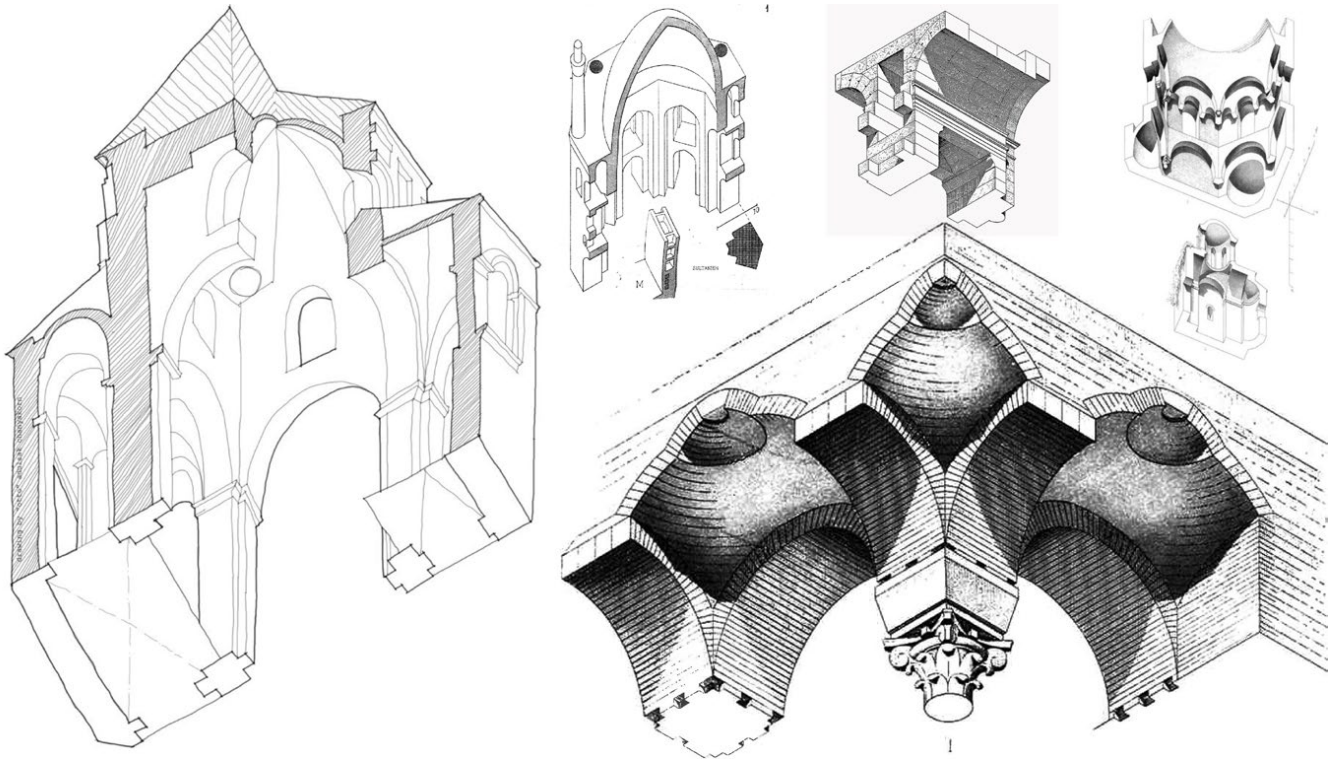
Half a century later, Le Corbusier would bring this idea to full expression through the concept of the promenade architecturale. For him, architecture was a spatial journey: an unfolding of forms, volumes, and views orchestrated for the moving body. Architecture was not merely built—it was curated in time.



Auguste Choisy, 1899 and Le Corbusier, 1937

It is in this lineage—from Piranesi to Choisy to Le Corbusier—that Yve-Alain Bois situates Eisenstein. Bois draws a conceptual genealogy across three centuries: Piranesi (18th), Choisy (19th), and Le Corbusier (20th), with Eisenstein standing not just as a cinematic figure but as a critical theorist of spatial experience.

A close reading of both volumes of Auguste Choisy's Histoire de l'Architecture reveals a consistent visual language: nearly all buildings are rendered through axonometric projections. Yet, one notable exception stands out—his representation of the Acropolis of Athens, for which Choisy abandoned axonometry in favor of a perspectival storyboard. Sergei Eisenstein recognized the significance of this rare deviation. It was not a stylistic indulgence but a critical decision—a recognition that the experience of space, particularly in the case of the Acropolis, could not be adequately conveyed through orthographic means alone.



Drawings from Histoire de l'Architecture (1899), Auguste Choisy

Choisy's use of perspectival sequencing suggested that the ancient Greeks conceived their architecture not as a static composition, but as a spatial narrative—a deliberate orchestration of visual moments experienced in motion. In contrast, the architects of the 19th-century Beaux-Arts tradition often sought symmetry and compositional unity, sometimes even disguising asymmetry through representational conventions. Choisy diverged from this tradition. While he was not the first to suggest that movement shaped spatial experience—figures like Viollet-le-Duc had already used theatrical metaphors such as mise en scène and pondération des masses—Choisy was arguably the first to systematically study how spatial sequences could shape perception.

This attention to perception resonated deeply with Eisenstein. He drew a parallel between Choisy's architectural phenomenology and his own theory of cinematic montage: in both, sequences—whether of buildings or shots—are structured to guide and manipulate the viewer's experience.

Choisy's usual tool for such spatial expression was axonometric drawing. In his own words:

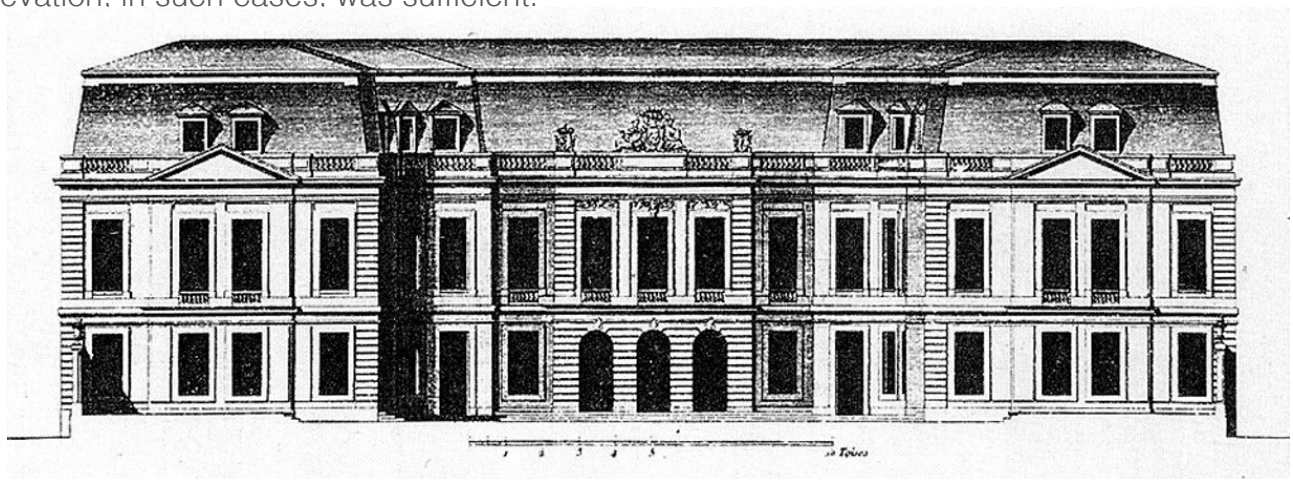
> *"In the system of axonometry, a single image, agitated and animated like the building itself, replaces the abstract figuration fractioned in plan, section, and elevation. The reader has in front of their eyes, simultaneously, the ground plan, the exterior of the building, its section and its interior disposition."*

For Choisy, axonometry offered more than the sum of plan, section, and elevation—it unified multiple viewpoints into a single, perceptually rich framework. Just as Eisenstein understood that paintings could generate virtual movement, Choisy saw axonometry as capable of activating imagined movement through space. The lack of a fixed perspectival viewpoint allowed the viewer's eye to "graze" across the drawing, assembling a coherent mental image of spatial organization and internal dynamics. Axonometry became a vehicle for what we might call a cinematic reading of architecture.

This stood in contrast to the preference of other 19th-century historians like Viollet-le-Duc, who often favored fixed-perspective renderings. Such images, while precise, limit the viewer's interpretive mobility—they enforce a singular vantage point, thereby curtailing the imaginative possibilities of spatial comprehension. Though one might assemble a sequence

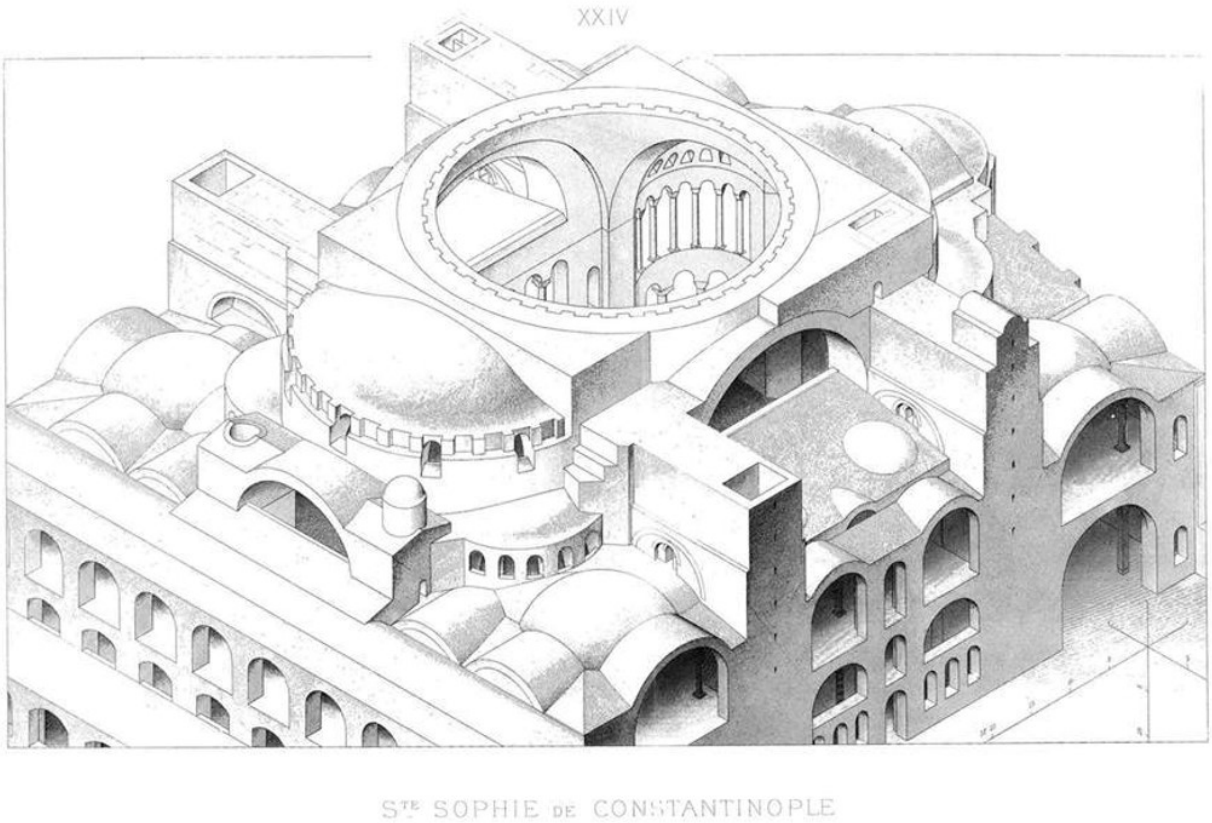
of perspectival views to simulate movement, this was impractical in the context of Choisy's encyclopedic project, which already spanned hundreds of drawings. Axonometry, by contrast, provided a compact yet dynamic system for articulating architectural experience.

Choisy's axonometrics suggested how buildings shifted in appearance depending on the viewer's location, allowing a kind of cognitive montage. The eye could scan across the image; the mind could synthesize perspectives, sectional logics, and volumetric comprehension. The absence of a central vanishing point allowed for permutability—an architecture of visual plurality, aligned with the phenomenology of movement.
Still, Choisy was selective. In cases where he judged interior space to be architecturally unremarkable—such as in many French hôtels particuliers—he defaulted to flat elevations. These buildings, for Choisy, were primarily façadic; their interest resided on the exterior. Elevation, in such cases, was sufficient.



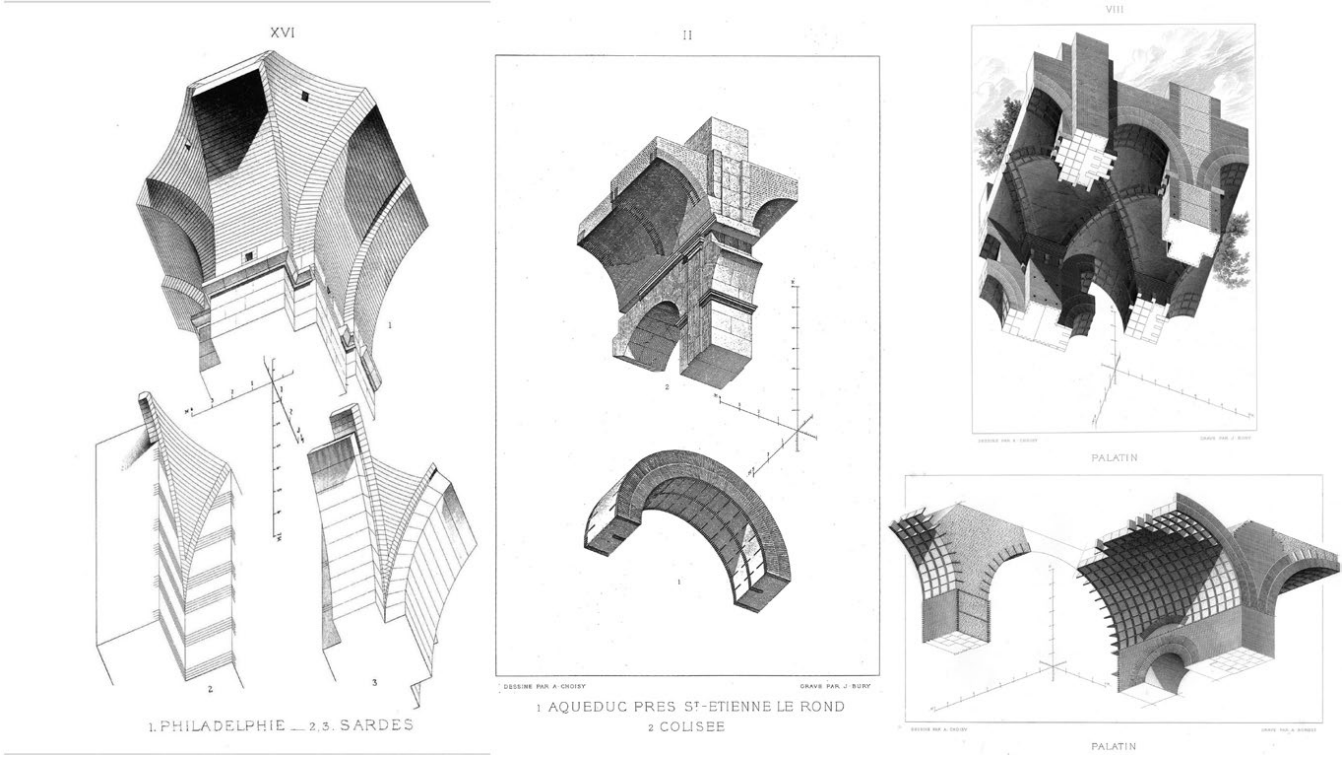
Drawing from Histoire de l'Architecture (1899), Auguste Choisy

This judgment echoes Eisenstein's own stance: some buildings are more cinematic than others. Unsurprisingly, both Choisy and Eisenstein gravitated toward similar architectural subjects—Hagia Sophia, for instance, appears in both their canons of study.



Drawing from Histoire de l'Architecture (1899), Auguste Choisy

Despite this resonance, a key difference underlies their motivations. Choisy, while attentive to shifting perception, remained committed to architecture as a product of tectonic logic. His rationalist lens prioritized structural clarity, spatial coherence, and technical innovation. It is for this reason that Baroque architecture is largely absent from his volumes. Choisy viewed the Baroque not only as theatrically excessive, but also as a deceptive mask—an aesthetic that concealed rather than revealed constructional truth. The Baroque, in his view, contributed nothing new to the tectonic imagination, and thus merited little discussion.



Drawings from Histoire de l'Architecture (1899), Auguste Choisy

This omission was not unique. It echoed earlier tendencies by figures like Piranesi and would later be mirrored by Le Corbusier. Rationalists across centuries often sidelined the Baroque, perceiving its ornamental excesses as antithetical to architectural integrity.

Eisenstein, however, did not share Choisy's rationalist priorities. His relative neglect of Baroque space had less to do with construction and more to do with narrative emphasis. Although his essay Montage and Architecture centers on Bernini's Baldacchino, Eisenstein largely avoids discussing its architectural form or the spatial grandeur of St. Peter's Basilica, which houses it. Yve-Alain Bois notes this omission as a curious hesitation. Instead of engaging with spatial sequencing, Eisenstein shifts his attention to symbolism and iconography, treating the Baldacchino not as space but as image.

St. Peter's Baldachin,
Gian Lorenzo Bernini

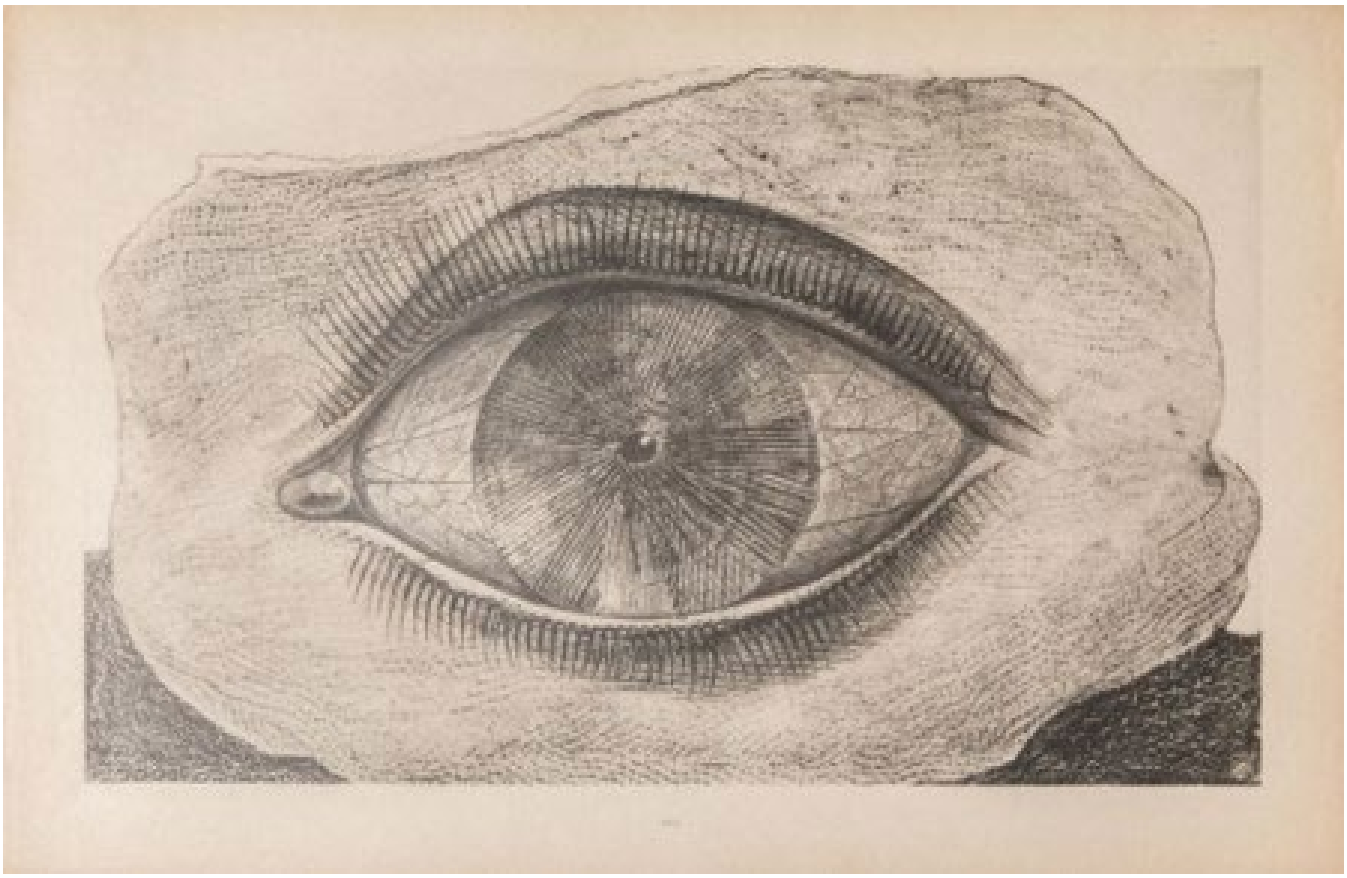
Thus, Bois identifies a shared silence between Choisy and Eisenstein regarding Baroque architecture—but one grounded in very different logics. Choisy dismissed it on the basis of tectonic inauthenticity; Eisenstein, despite his affinity for theatricality, sidestepped its spatial logic in favor of semiotic readings. In both cases, the Baroque remains a kind of unspoken excess—too ornamental, too opaque, or too discontinuous with the theoretical frameworks they were each constructing.

■ ■ ■

Notes on Eisenstein's text

Eisenstein begins his essay by reflecting on the concept of the path, a term he uses to describe both the trajectory of visual perception and the mental engagement that accompanies one's encounter with images and spaces. In cinema, this path is imaginary—it corresponds to the movement of the eye across the frame, shaped by compositional cues and the sequence of shots. In architecture, by contrast, the path is literal: it is the actual movement of the spectator through space, generating a series of unfolding views and spatial impressions.

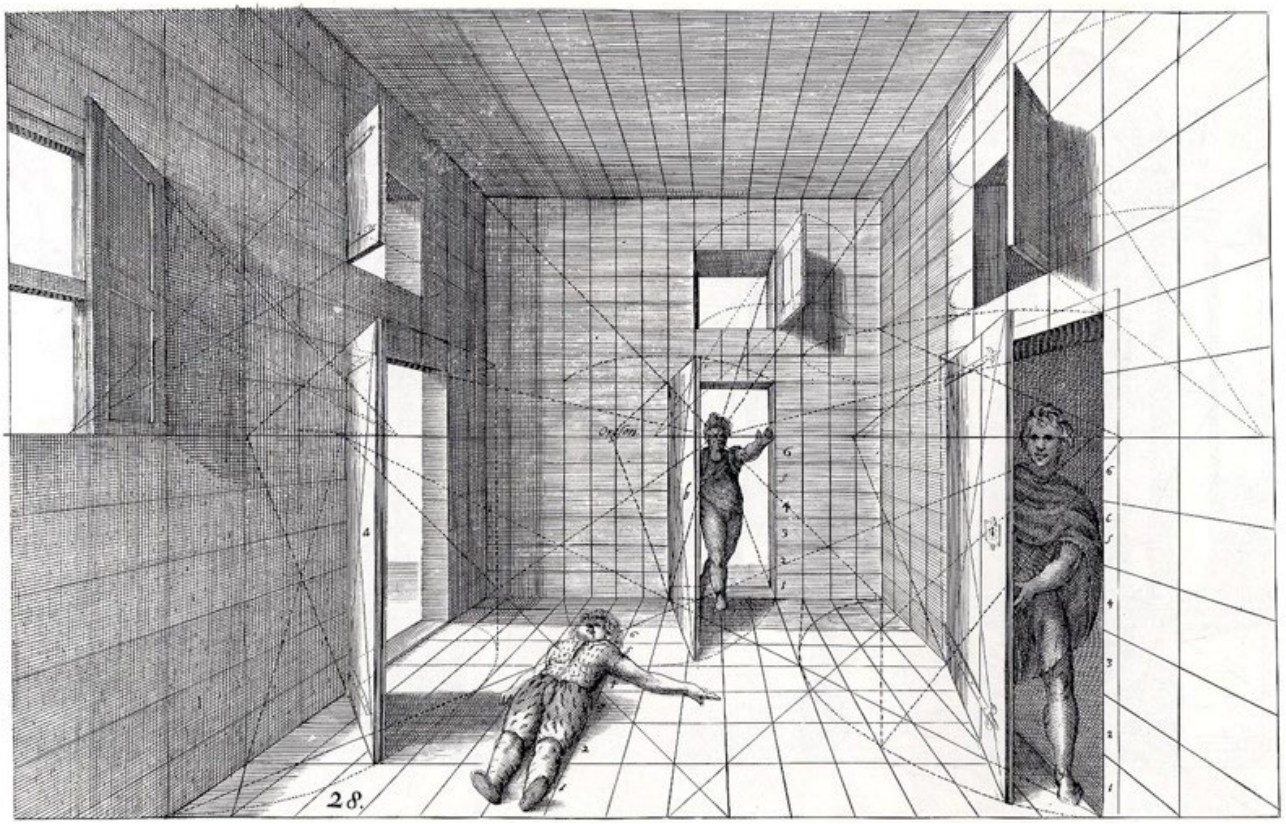
For Eisenstein, montage is not merely the technical act of cutting and splicing film; it is the very structuring of experience in time. Even within a single, static image, the eye traverses various focal points, creating a dynamic montage in the mind. This principle, he argues, extends beyond film to literature and architecture—any form where disparate elements are arranged to produce a coherent experiential sequence.



The wheel of light (1925), Max Ernst

What distinguishes architecture from cinema, however, is the inversion of this perceptual relationship. While film presents a succession of images to a stationary viewer, architecture positions the viewer as a moving agent within a relatively fixed environment. Yet the outcome is analogous: both forms orchestrate perception across time. For Eisenstein, this temporal sequencing of spatial encounters renders architecture inherently cinematic. Like film, it choreographs a path—physical or optical—through which meaning is gradually constructed.

Audience In The Newsreel Theatre (1940), Gene Lester



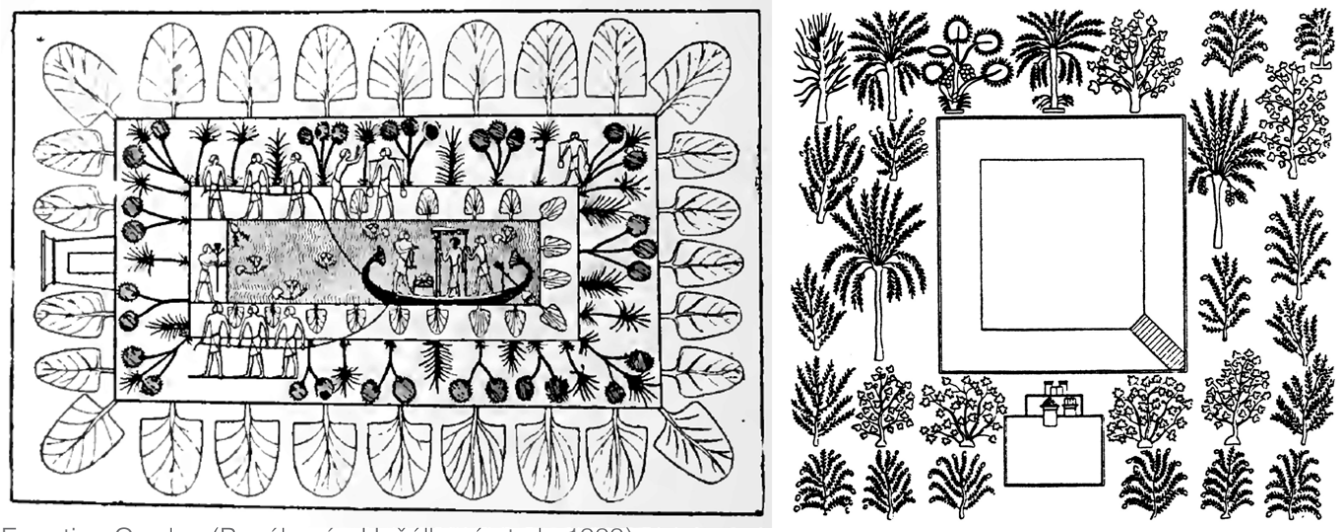
Perspectival scene (1604), Jan Vredeman de Vries

The sequential nature of perception, Eisenstein suggests, is both intuitive and universal. He illustrates this with children's drawings, where the absence of formal training in perspective allows for a raw, experiential rendering of space. Rather than fixing a single viewpoint, children often combine multiple perspectives within a single image, capturing not a static scene but the dynamic unfolding of spatial experience. Their drawings become visual records of movement—of their own shifting positions in space, and of their perceptual encounters over time. A child might, for instance, depict a row of trees in full elevation while rendering the path beneath them in plan view. Such hybrid imagery resists the constraints of conventional perspective and instead evokes the continuity of lived experience.

Recreation of a child's drawing by Sergei Eisenstein

For Eisenstein, these intuitive compositions reveal that the principles of montage predate cinema and exist at the core of human visual storytelling. Even in their most rudimentary creative expressions, humans organize space and time sequentially. The child's drawing does not freeze a single moment, like a photograph or perspectival sketch; it accumulates impressions into a continuous, composite whole. This peripatetic logic—of viewing as movement—mirrors how architecture is encountered in real life: not all at once, but progressively, through time.

Eisenstein finds similar strategies of sequential representation in ancient Egyptian painting. In one example, a pond is rendered in plan, while the surrounding buildings and trees are shown in elevation. The image does not obey perspectival unity; instead, it organizes spatial elements along an imagined path of movement. This results in a hybrid visual field, where the viewer mentally traverses the scene, reading it as a spatial narrative. Like montage in cinema, the composition guides perception temporally, through the juxtaposition of distinct yet related viewpoints.

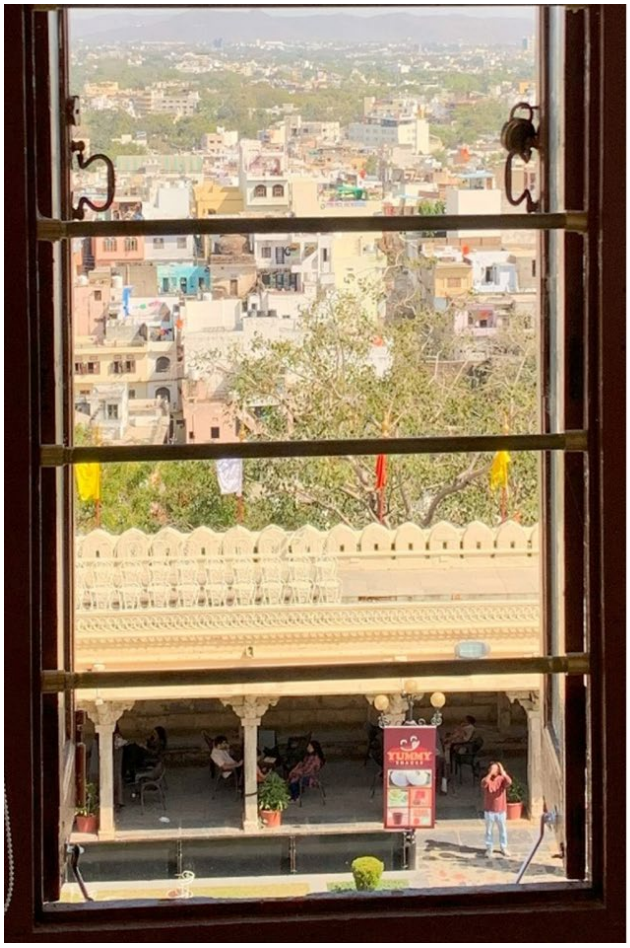


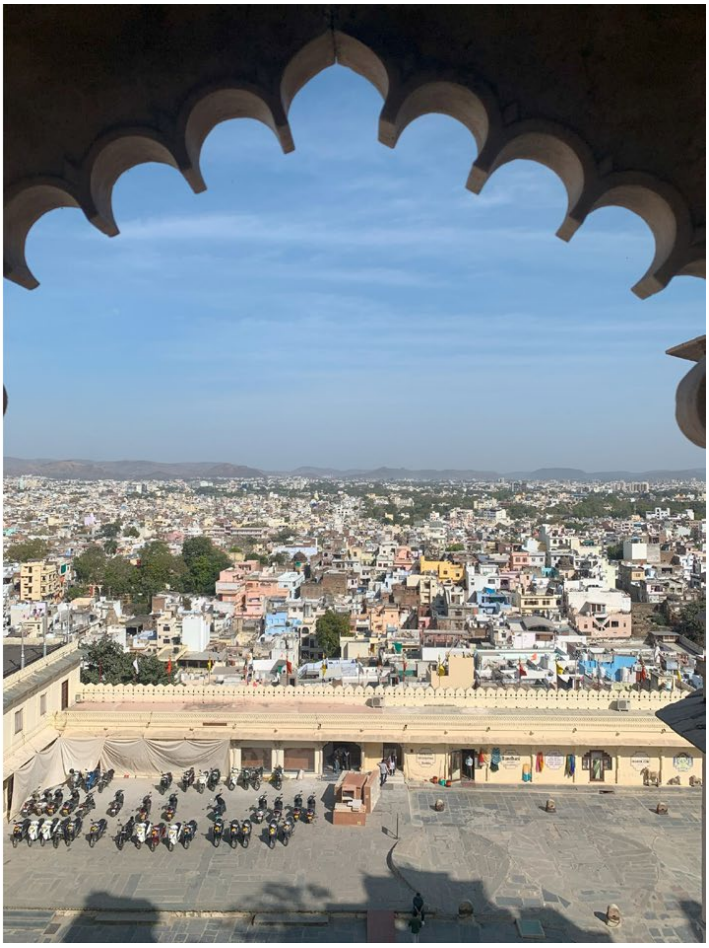
Egyptian Garden (Pacáková - Hošálková et al., 1999)

This mode of representation finds a striking parallel in the vernacular miniature paintings of Rajasthan. These artworks, often narrative in nature, not only depict unfolding events but also embed them within richly articulated architectural settings. Domed pavilions (chhatris), pillared halls (mandapas), jharokhas (overhanging balconies), terraces, courtyards, and verandahs—elements characteristic of Rajasthani architecture—are not merely scenic backdrops but active participants in the narrative structure. As Yve-Alain Bois notes of Eisenstein's cinematic staging, architecture here functions not as passive scenery but as an "active agent of the plot." The depicted spaces are both enabling and expressive; they provide the spatial conditions for the events to occur and, at the same time, guide the viewer's visual journey through the scene.

Maharana Ari Singh II enjoying Jagmandir (1767), Attributed to Jiva and others

Indian miniature paintings frequently defy the conventions of single-point perspective, instead combining frontal, planimetric, and oblique projections into a single frame. This multiplicity of views not only allows for a fuller depiction of architectural context but also reinforces the temporality of perception. Even from the windows of the Udaipur City Palace, one can experience a similar visual phenomenon: due to its elevated siting on a ridgeline overlooking the lake and city, the view reveals a layered, semi-elevational panorama—part plan, part elevation. The eye scans across the urban fabric as though reading a sequence of shots, framed by the palace window itself. The city becomes cinematic, and the window, its lens.





View of the city from windows of Udaipur city palace

In the early twentieth century, as the so-called decadence of representational art set in, a paradoxical return to archaic visual strategies began to emerge. For Eisenstein, this regression was not a sign of decline but a premonition of cinema. The reappearance of primitive techniques—fragmentation, sequentiality, and multiperspectivity—signaled a subconscious anticipation of cinematic logic. Artists, with or without consciously intending it, began to re-engage with compositional methods that echoed the principles of montage.

Abstract landscape, David Burluik
Eisenstein's decomopsition and

In the works of David Burliuk and Robert Delaunay, Eisenstein observed this revival of non-perspectival visual language. Burliuk's paintings, with their fractured forms and juxtaposed visual planes, enact a form of perceptual dislocation: vision is not unified but sequenced, encouraging a temporal reading of space. Likewise, Delaunay's dynamic renderings of the Eiffel Tower break the object into angular fragments and overlapping planes. The viewer encounters multiple perspectives simultaneously, suggesting not the fixity of form but its unfolding—almost as if motion were embedded within the static image.

Paintings of Eiffel tower by Robert Delaunay (1909,1910,1885)

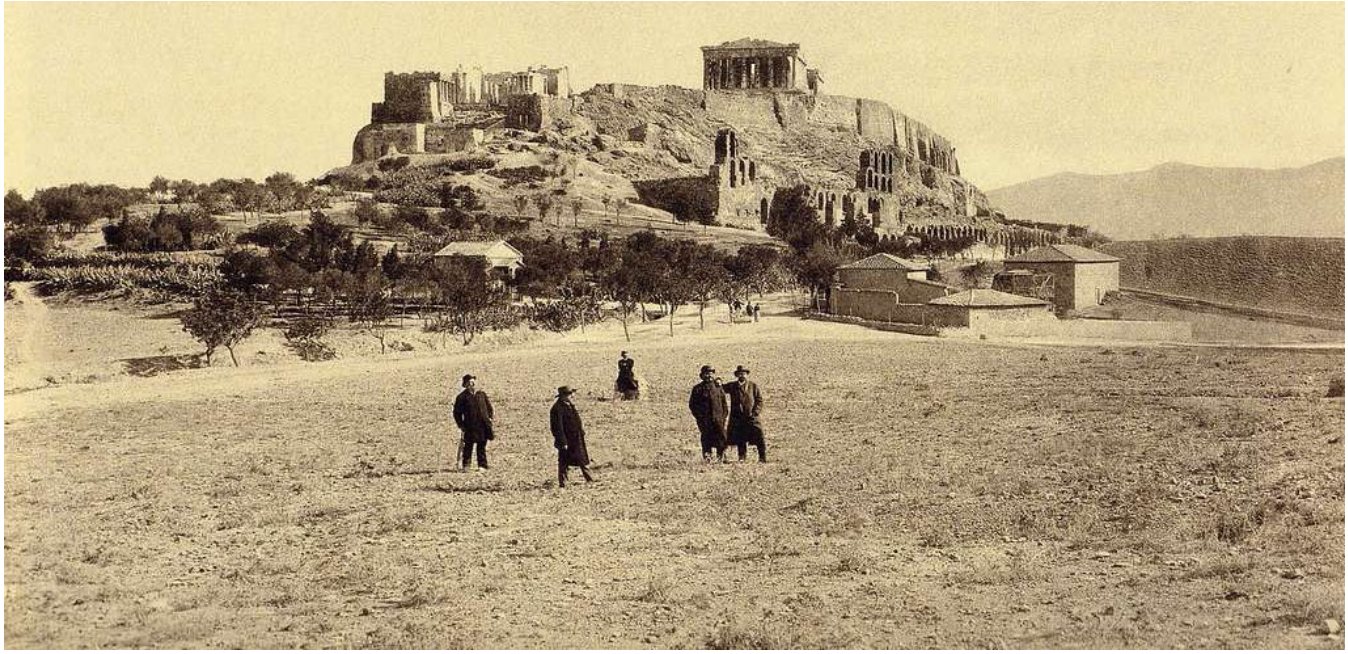
Eisenstein's interpretation of this artistic transformation is less concerned with historical accuracy than with conceptual insight. He proposes that modern art, having exhausted its mimetic engagement with external reality, begins to turn inward. Just as early visual art was grounded in internal perception rather than realistic depiction, so too does modernism retreat from outward representation toward subjective experience. The visual field becomes less a window onto the world than a projection of inner vision. This shift dismantles spatial coherence and perspectival unity, replacing them with abstraction and expressive fragmentation.

This "regression," for Eisenstein, is not a decline but a necessary inversion—a retraction into interiority that sets the stage for cinema. The breakdown of spatial logic in painting becomes the ground from which montage, as a temporal-spatial logic of perception, can emerge. In this sense, montage reintroduces not only time and space into visual experience but also a kinetic engagement with perception itself.

Pablo Picasso's work :
Science and Charity (1897)
La Vie (1903),
Au Lapin Agil (1905),
Les Demoiselles d'Avignon (1907),
Femme assise (1909),
Femme assise dans un fauteuil (1913),
Guernica (1937)

Despite modern painting's experimentation with distortion, abstraction, and spatial disruption, Eisenstein argued that its inherent limitation remained: it was static. No matter how radical its transformations, visual art could not fully overcome its temporal immobility. The totality of perception—the multidimensional unfolding of space over time—remained out of reach. It is here that cinema offers a radical departure. Through movement and sequence, film enables a temporal and spatial synthesis that painting could only approximate.

Eisenstein credits this technological and perceptual breakthrough to the invention of the Kinetograph by William K. L. Dickson. The camera's capacity to record motion allowed visual representation to transcend the confines of the static image. At last, the "spirit" of visual experience could be liberated from the fixed surface, and cinematic montage could materialize what earlier visual forms had only gestured toward.
Eisenstein intriguingly suggests that architecture—specifically, ancient Greek architecture—served as a precursor to cinema. In the Acropolis, he identifies a raw and elemental cinematic quality: the organization of space as a sequential visual narrative. Just as a film unfolds through the movement of images over time, architecture unfolds through the bodily movement of a spectator in space. Drawing a parallel to medieval cathedrals, Eisenstein refers to them as "books in stone," implying that buildings, like films, can tell stories—only their language is spatial rather than temporal.



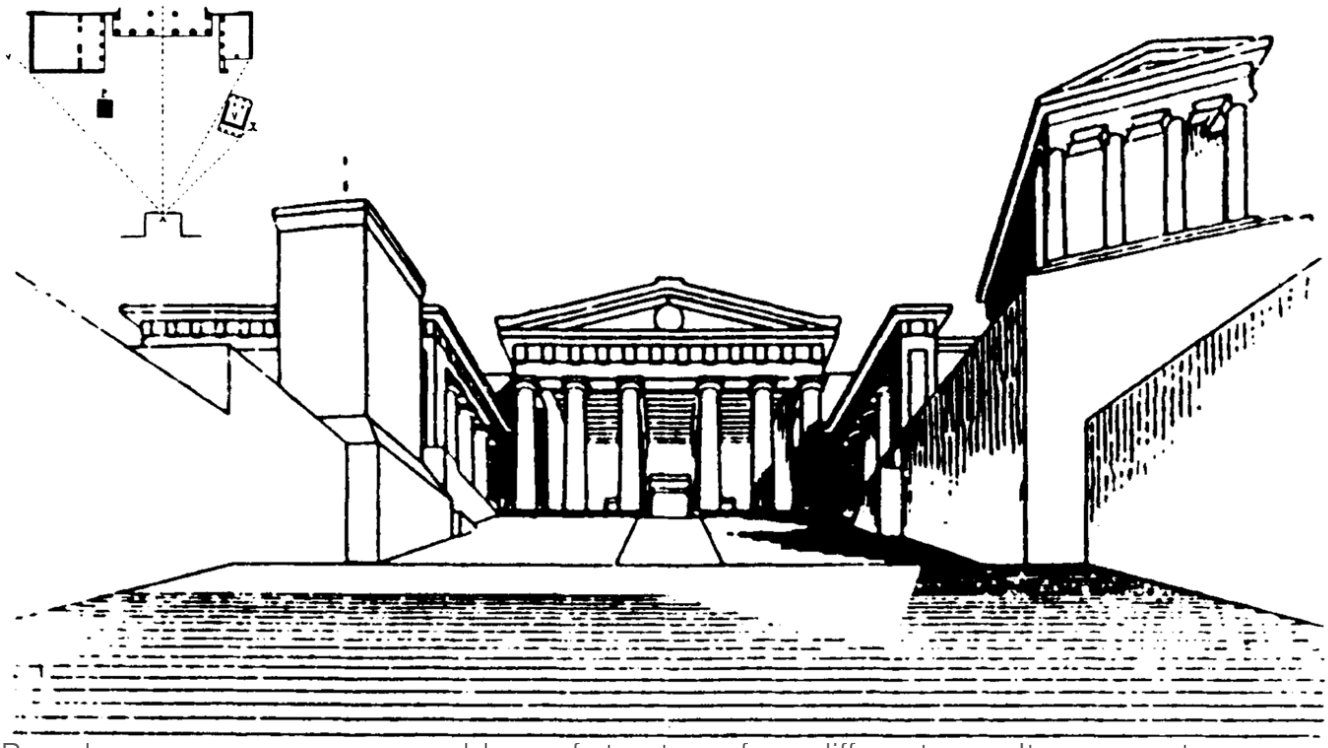
Acropolis in late 19th century

At this pivotal point in his essay, Eisenstein offers a striking formulation:

> *"It is hard to imagine a montage sequence for an architectural ensemble more subtly composed, shot by shot, than the one that our legs create by walking among the buildings of the Acropolis."*

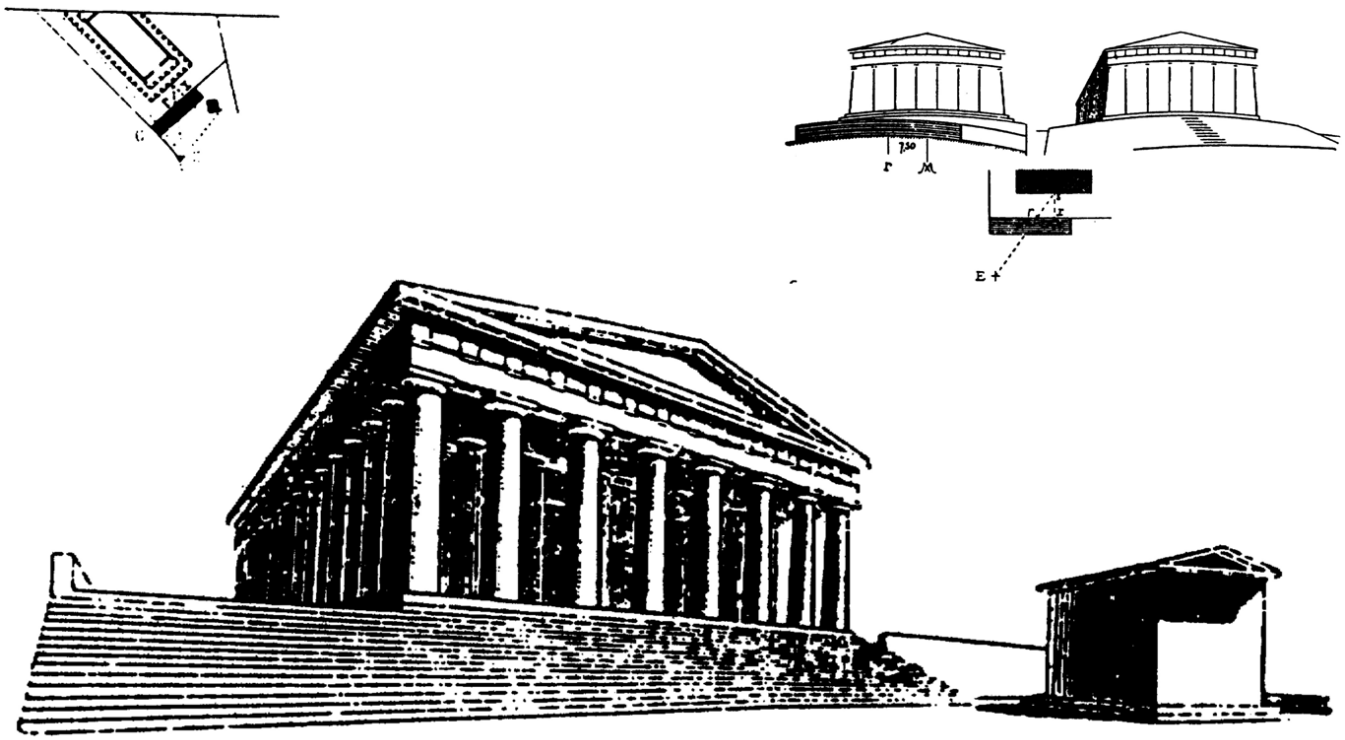
This observation draws directly on Auguste Choisy's analysis in Histoire de l'architecture, which Eisenstein uses to illustrate how architectural composition can be interpreted in cinematic terms. The Acropolis, he argues, offers a kind of "architectural montage," whose experiential sequence—as one moves through and between its structures—functions analogously to film editing, guiding perception and structuring meaning through spatial progression. However, the lengthy passage Eisenstein quotes from Choisy—though precise and rich in detail—can be challenging for readers unfamiliar with the ruins of the Acropolis or the specific architectural elements referenced by their historical Greek names. Without visual context, the architectural sequence may feel abstract or difficult to grasp.

It is therefore helpful to recall Choisy's perspectival storyboard and the key spatial moments Eisenstein draws from it. The sequence, as laid out in the storyboard, unfolds as follows:

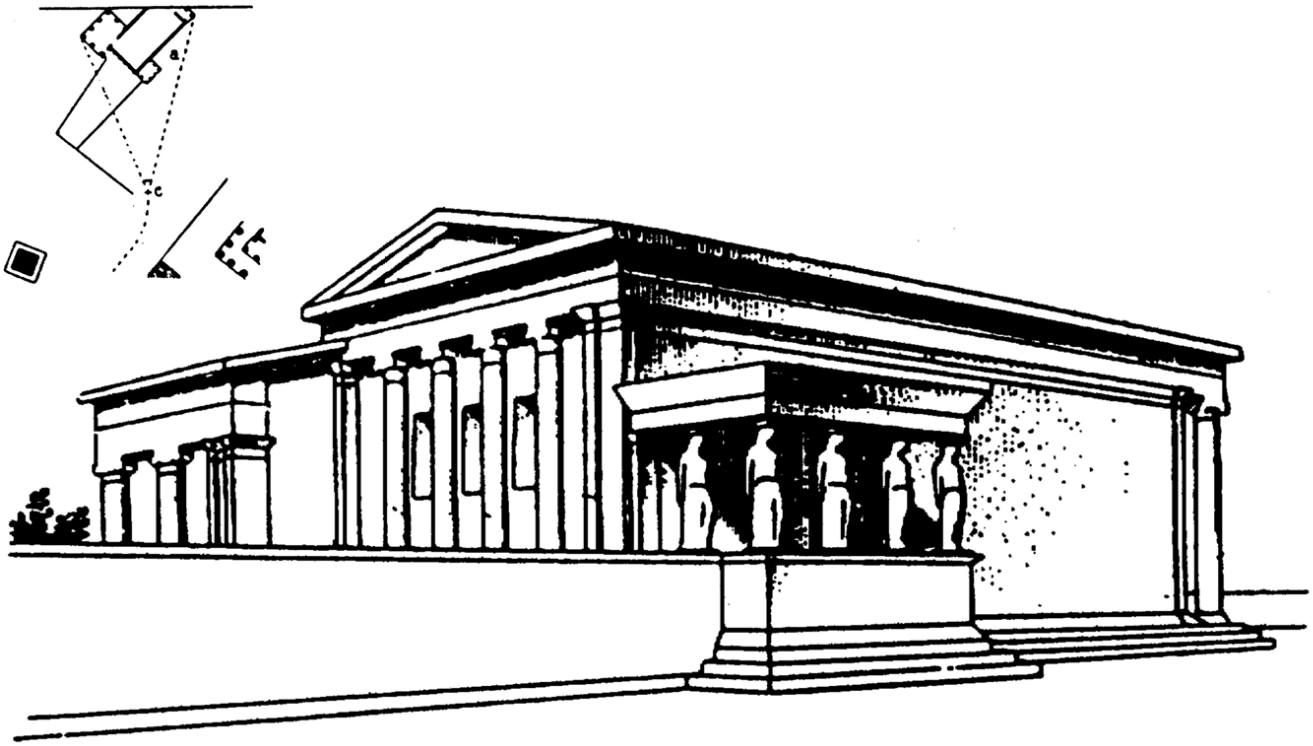


Propylaea appears as an assemblage of structures from different eras. Its asymmetry, as Choisy notes, contributes to its "picturesqueness." He emphasizes this with a plan, remarking that "nothing could be more uneven than the plan."

The next view reveals the towering figure of Athena Promachos, partially obscuring the Erechtheion and the Parthenon in an oblique background composition.



The third moment in the sequence offers a diagonal, almost cinematic view of the Parthenon itself.



Finally, the sequence culminates with a frontal view of the Erechtheion, its Porch of the Caryatids dominating the field of vision.

This perspectival choreography, as interpreted by Eisenstein, mirrors the structure of montage in cinema: each architectural encounter builds upon the last, revealing new spatial relationships and reorienting perception. Because the original descriptive passage is densely packed with references, this interpretive retelling—with visual support—helps clarify how Choisy's architectural analysis becomes the very ground on which Eisenstein constructs his theory of spatial montage.

Eisenstein once defined cinematism as the conceptual medium between painting and cinema. Extending this lineage, this essay proposes a third medium—simulationism—positioned between cinema and architecture. If cinema choreographs vision through time, and architecture unfolds in physical movement through space, simulationism operates in algorithmically constructed virtual space—where movement is at once embodied and imagined.

As the digital reconfigures both time and space, we find ourselves in what might be called the era of neo-simulationism—a new stage in the dialectic of spatial aesthetics. Following Hegel's triad: painting as thesis (art in static virtual space), cinema as antithesis (art in dynamic virtual space), and simulation as synthesis (art in real-time, reconstructed space), this new condition echoes Henri Lefebvre's notion of produced space—no longer passively perceived, but actively authored through interaction. What we see is not a representation of space, but a space of representation, layered with symbolic, bodily, and algorithmic agency.

Where cinema fixed the viewer's gaze within a linear montage, simulation opens the experience to multiplicity—what Gilles Deleuze might call a movement-image unbound from chronology, allowing rhizomatic pathways through digitally rendered worlds. And if Walter Benjamin described architecture as absorbed in a state of distraction—perceived in a state of peripheral attention—simulation now demands a new kind of immersion: a form of critical participation that merges presence with projection.

Simulationism fuses cinematic montage with built environments and interactivity, generating environments that begin to script themselves as they are traversed. This recalls Robert Delaunay's concept of simultaneism, in which spatial and temporal fragments are collapsed into a unified visual field. But here, rather than being manually composed like painted fragments, the construction is algorithmic—infinitely recomposable and responsive in real time.

Virtual reality, like cinema, is a medium of virtual perception: it unfolds along an imagined path shaped by the movement of the eye, even while the body remains still. Yet unlike cinema, this path is not predetermined by an editor or director. It is choreographed—improvised, even—by the user. In this respect, simulation mirrors architecture: a space one navigates at will, assembling experience through movement. But unlike architecture, simulation does not unfold in the physical world. It occupies a synthetic terrain: perceptually immersive, but materially dislocated.

This essay proposes to use interactive virtual simulation as a theoretical instrument to re-examine the spatial logic of montage. More precisely, it turns to the reconstruction of the Acropolis within Assassin's Creed Odyssey, conceptualized and developed by Ubisoft. The aim is not aesthetic validation but pedagogical inquiry: to investigate how digital reconstruction can elucidate the theories of spatial sequencing found in the writings of Auguste Choisy and Sergei Eisenstein. In this sense, the simulation is not the subject of critique, but the lens—a medium through which historical spatial thought may be reanimated and re-experienced.

(UbiSoft)

The word Acropolis derives from two Greek roots—akron (high) and polis (city)—and thus translates, quite literally, to “high city.” Perched atop the most elevated terrain in Athens, the Acropolis stood as both a visual and symbolic apex of the city’s religious and civic life.

Yet the spatial experience of the Acropolis does not begin at its formal entry. It begins long before, with the first distant glimpse of the citadel as one enters the city. Even this early moment activates a sequence of perceptual transitions—a visual prelude in which the monument first appears as a silhouette on the skyline, remote and elevated, suspended above the rhythms of daily urban life.

As one moves closer through the dense weave of Athens, the topography begins to stage a spatial drama. The steep terrain gradually conceals what was once visible. The Acropolis disappears—not through architectural framing, but through natural occlusion. The sacred hill, once prominent, vanishes behind the very city it presides over. This erasure is not incidental—it produces a delay, a tension, a pause in visual access that heightens the eventual reappearance.

The path of ascent becomes a cinematic device, even before any monument is formally encountered. It organizes space into a choreography of concealment and anticipation. The sacred does not reveal itself immediately. It is withheld—not absent, but latent—waiting for the right frame, the right distance, the right gaze.

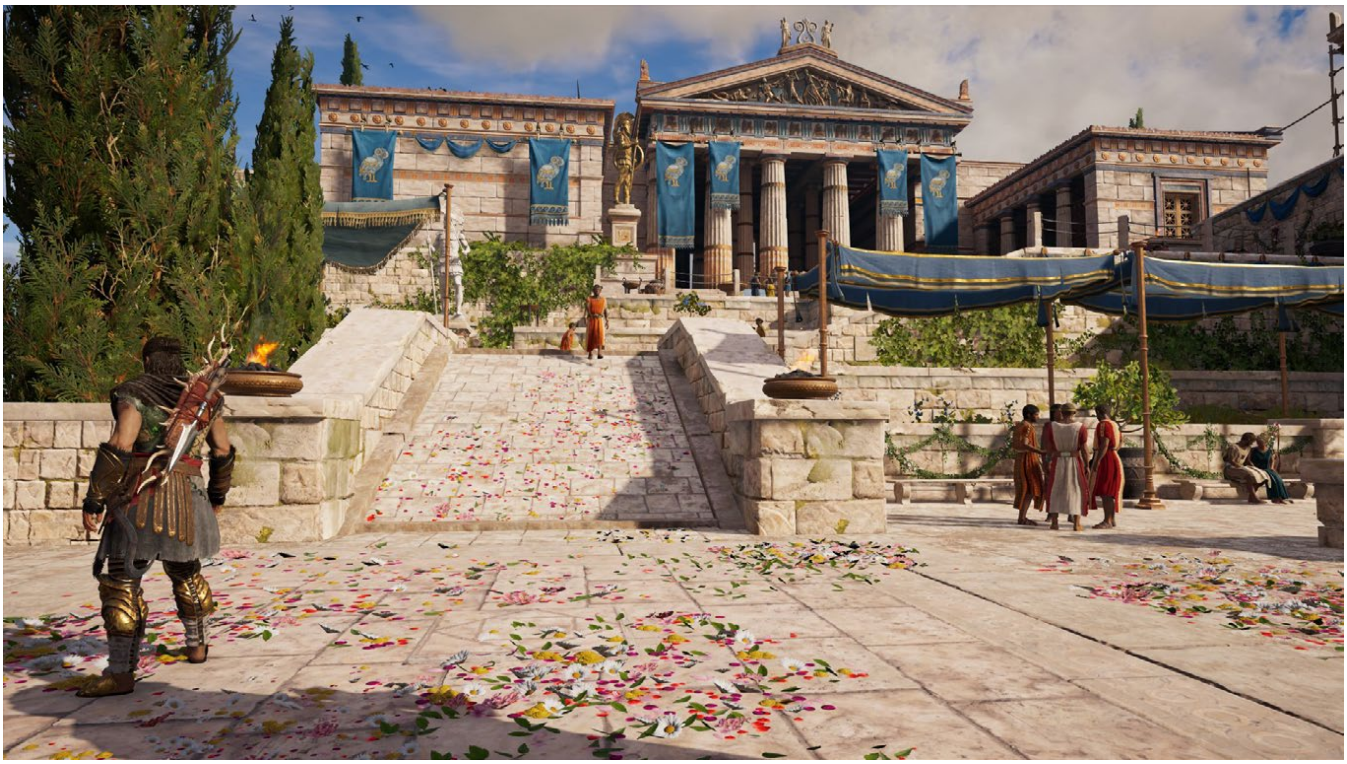
View of Acropolis far from distance

The climb to the Acropolis is not merely physical—it is symbolic. As the visitor ascends, they are gradually withdrawn from the chaos of the city below, entering into a state of perceptual and psychological isolation. This detachment prepares them for the sanctified realm above.

Disappearance of Acropolis

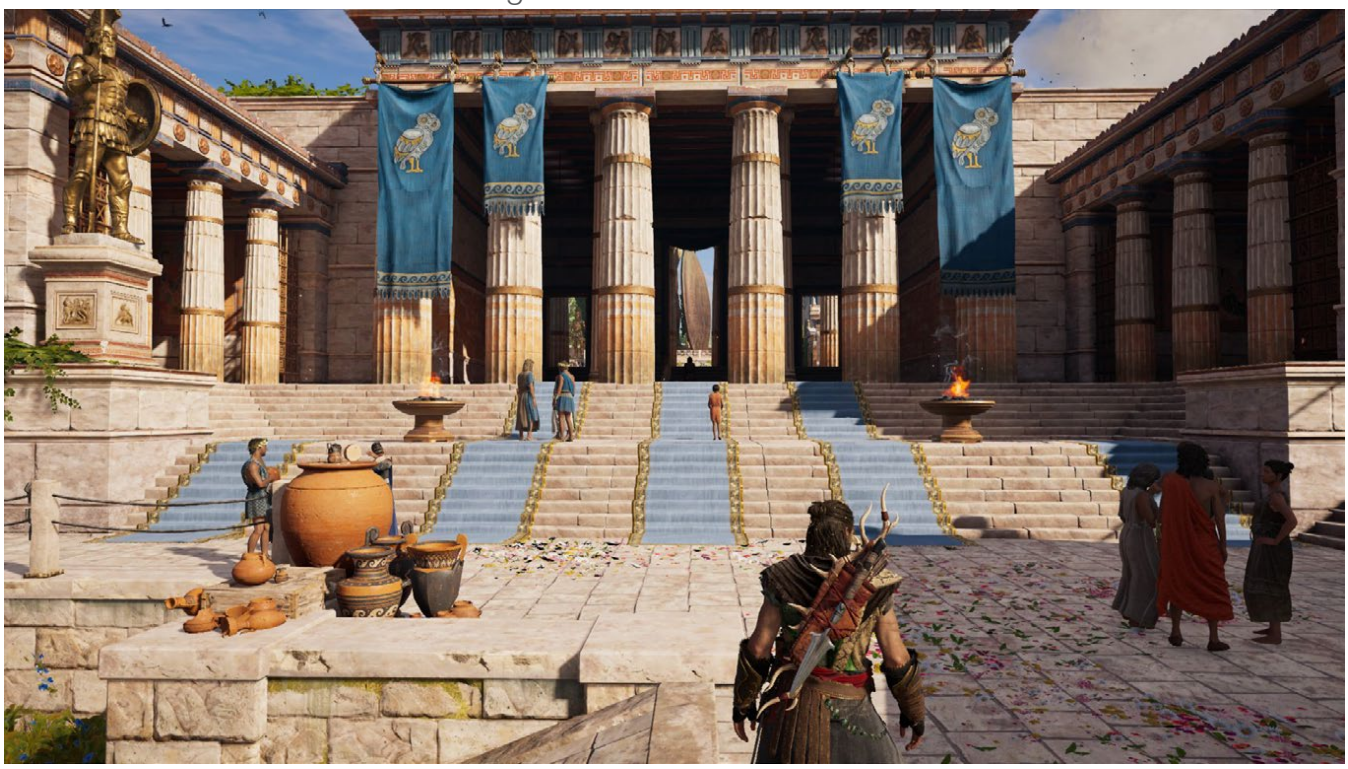
At the summit, the Propylaea appears—not simply as an entrance, but as a threshold. It is the first architectural frame in Choisy's storyboard, and thus the first shot in the cinematic unfolding of the Acropolis experience. Conceived as a monumental gateway, the Propylaea accommodated the vast crowds of the Panathenaic procession, a ritual that culminated here every four years. While its plan may appear irregular—an asymmetry dictated by the sloping terrain and topographical constraints—its massing, as Choisy emphasized, achieves a compositional coherence that is 'picturesque' rather than classically symmetrical.

Importantly, the Propylaea is not a flat facade, not a simple passage. It operates as a volumetric transition, mediating between the civic world and the sacred precinct. It does not only mark a spatial boundary—it creates a temporal delay, ushering the visitor through a zone of dimmed light and narrowing perspective, preparing them for the revelation beyond.



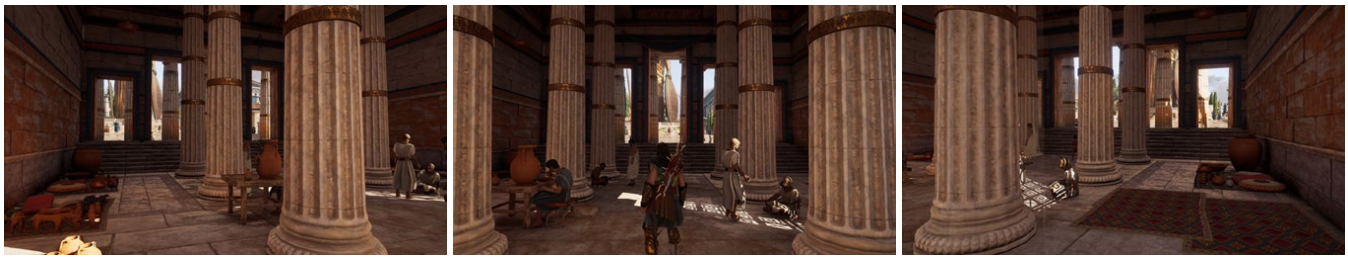
View of Propylaea (entrance gateway)

As one approaches this gateway, glimpses begin to emerge. Framed in the darkened interior, and partially concealed, is the glint of bronze—Athena Promachos. The sculptural form is not revealed in its entirety; it is withheld, partially masked by the thick architectural envelope of the gateway. This contrast between darkness and the gleaming metal intensifies the perceptual drama. Light and shadow here are not incidental—they are compositional elements in the architectural montage.



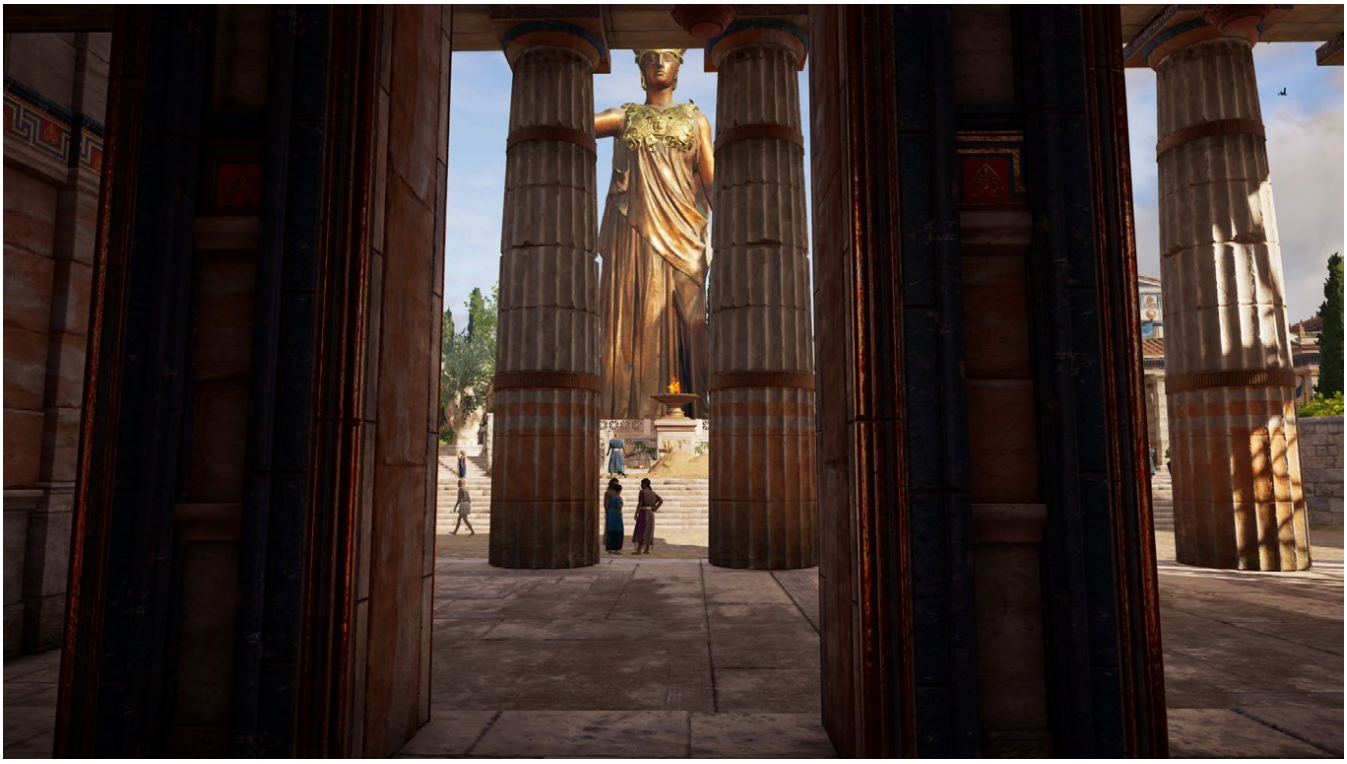
Fragmented framing of Athena Promachos (Bronze statue of Athena)

Once inside the Propylaea, the viewer's perception fragments. The path no longer runs along a singular axis, and the statue remains visually incomplete. Only portions of the statue are visible through the side openings. The full form of the sculpture cannot yet be grasped. The viewer is compelled to continue moving—to edit together the space in time, frame by frame, as in montage.



Perceptual variations inside Propylaea

As the visitor approaches the exit of the Propylaea, the colossal bronze statue of Athena—Athena Promachos—begins to emerge. This is a moment of parallax: though the viewer moves closer, the sheer scale of the figure causes it to recede visually, appearing to shrink even as it looms larger in space. This inversion of perception—where spatial proximity does not equate to visual dominance—is a key element in the architectural montage.



Parallax effect : statue in perception becomes smaller as the envelope of Propylaea decreases

Upon exiting the gateway, Athena Promachos is finally revealed in full—a monumental figure now fully commanding the frame. This is the second shot in Choisy's perspectival storyboard. The statue, towering above, dominates the field of vision, obscuring much of what lies behind. Only a narrow oblique glimpse of the Parthenon appears in the distance, slightly displaced to the right. As Choisy observes, this visual panorama is structured entirely around Athena Promachos; the other architectural elements recede into a supporting role, reinforcing the statue's centrality in the spatial composition.

Revealtion of Athena Promachos

It is only as the viewer continues forward, passing the statue, that the Parthenon begins to assert its presence. The shift in visual emphasis is not abrupt but choreographed—structured as a cinematic cut from one protagonist to another. And as one moves closer to pay homage at the statue's base, its monumental form becomes visually incomprehensible: it overwhelms the frame, loses its legibility as a whole, and becomes too close to be seen. What was once the focal point now exceeds the limits of perception.

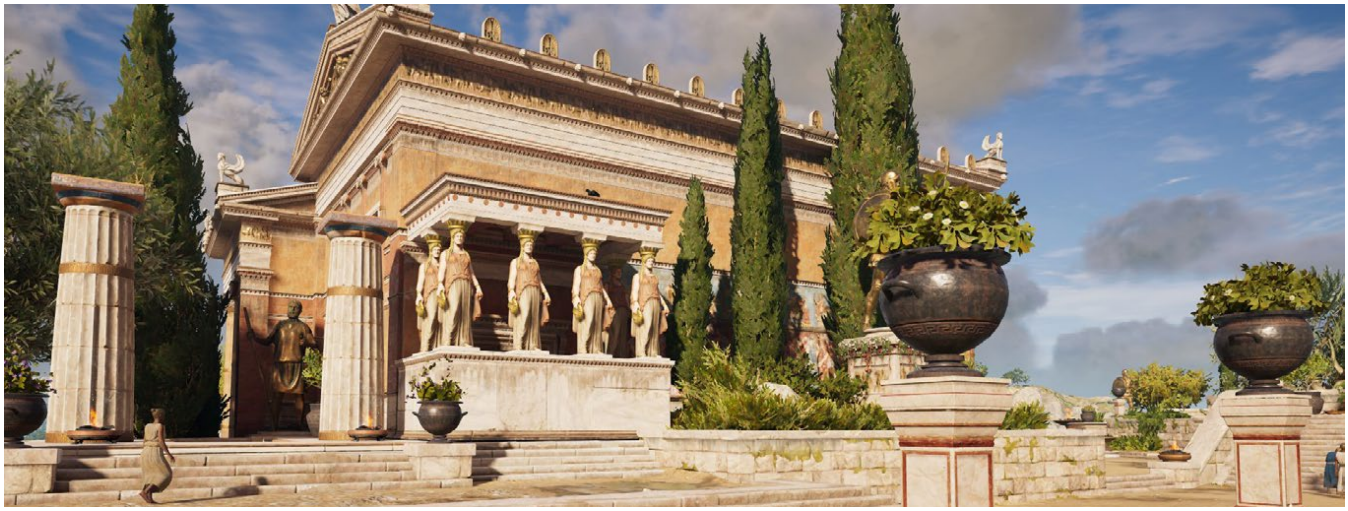
The statue becomes too large to comprehend fully in close proximity

Here, Choisy highlights how the siting of the Parthenon defies the prevailing architectural conventions of his own time—especially the expectation of axial symmetry. The most prominent monument of the Acropolis is not aligned directly with the Propylaea, the ceremonial gateway; instead, it is situated at an oblique angle. This disalignment is not accidental but intentional, reflecting an ancient Greek preference for a picturesque composition over rigid frontal order.

Oblique view of Parthenon

Rather than impose geometric regularity onto the terrain, the Greeks responded to the natural contours of the site, allowing the architecture to evolve organically. The oblique view, far from being a compromise, was a calculated choice. As Choisy notes, while frontal façades were sometimes used for monumental emphasis—as in the Propylaea and perhaps in the Pronaos of the Parthenon—these were exceptions rather than rules. Even in the case of the Parthenon, the first encounter is through an angled perspective, challenging any clear categorization. The oblique reveals the building slowly, withholding its full form in favor of a temporal unfolding. At this stage of the spatial journey, the Parthenon commands the field of vision. Yet as the viewer moves closer, proximity paradoxically renders the monument visually illegible. The architecture escapes full apprehension; its scale surpasses the frame of human vision. The gaze, unable to resolve the whole, begins to drift—drawn toward what had previously been concealed: the Erechtheion.

Erechtheion which was concealed behind the colossal bronze statue is now in sight



Porch of Caryatids stands out in relief as the minimal wall behind it acts as a backdrop

On the far side of the Acropolis plateau, the Erechtheion emerges as the next focal point in the sequence. Smaller in scale and more intricate in form, it now occupies the center of perceptual gravity. The famous Porch of the Caryatids comes into view, its sculpted columns delicately poised—a stark contrast to the monumentality of the Parthenon. This juxtaposition is not just formal but experiential: where the Parthenon overwhelms, the Erechtheion invites a more intimate engagement. As the silhouette of the Erechtheion emerges, its austere, planar wall serves as a neutral backdrop, allowing the Porch of the Caryatids to stand out in bold relief. The contrast enhances the sculptural presence of the figures, whose ornate detail is magnified by the simplicity that surrounds them.

The decision to incorporate such elaborately carved columns invites speculation. While historians continue to debate the identity and symbolic role of the Caryatids, from a spatial and narrative standpoint, one might interpret them as perpetual witnesses to the ritual climax of the Panathenaic procession. Their poised, silent gaze aligns with the path of arrival, suggesting not merely structural function but theatrical intent—as if they occupy the threshold between architectural frame and mythic stage. The virtual simulation—while remarkably detailed—fails to fully capture the individual distinctions between the Caryatids. In the original, each figure was sculpted with subtle variations: a tilt of the head, a shift in drapery, a difference in bearing. These differences contribute not only to the realism of the ensemble but to the sense of animated stillness, further amplifying their role as spatial narrators rather than passive ornament.

Maidens of the porch of Caryatids which looks over to the procession pathway

The Panathenaic procession culminates at the sacrificial altar, situated near the rear of the Erechtheion. Here, a stepped platform provided a stage for offerings made in honor of Athena. This final ritual act directs the movement toward the eastern porch of the Erechtheion—unlike the Caryatid porch, which was closed from the outside, this entrance is accessible. It leads to the Shrine of Athena Polias, which once housed the sacred olive-wood statue of the goddess. This figure, central to the cult of Athena, was ceremonially adorned during the festival and marked the spiritual terminus of the grand procession.

The sacrificial alter on the right side with steps of

Continuing onward, the viewer's path once again intersects with the oblique perspective of the Parthenon. Its visual dominance is reestablished—not through a direct confrontation, but through the slow, spatial return of its form into the spectator's field of vision. In this way, the Parthenon bookends the spatial experience, reasserting its symbolic and architectural authority.

Oblique view of Parthenon is introduced again as one moves towards the main entrance of Parthenon

At every major point in the journey, a single architectural element dominates the composition. Rather than a chaotic overlap of forms, the Acropolis is choreographed as a series of visual episodes: each structure steps into prominence, while the others recede as background or framing devices. This carefully orchestrated sequence—alternating between concealment and revelation, compression and expansion—mirrors the principles of montage, where meaning arises not from isolated images, but from their calibrated succession.

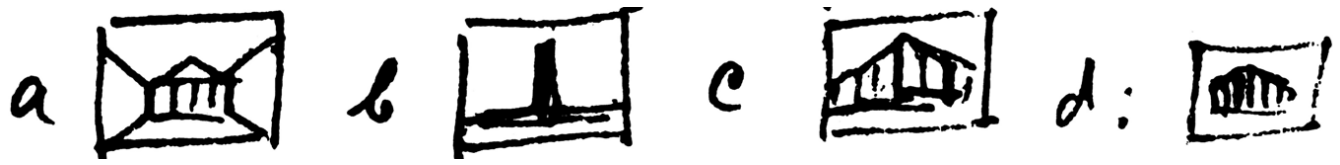
Whether consciously orchestrated or intuitively shaped, the spatial decisions made by the ancients play a vital role in the dynamic perception of the Acropolis. Choisy observes that while Athena Promachos dominates the scene, her monumental base simultaneously conceals the Porch of the Caryatids, preventing its delicate grace from being overwhelmed by her scale. Only when the viewer passes the statue and moves into closer proximity does the porch emerge—preserved in its visual autonomy and protected from monumental overshadowing. At such close range, Athena Promachos herself can no longer be perceived in totality, reinforcing a choreography of fragmented revelations.

Choisy emphasizes the calculated nature of first impressions at each moment of encounter. The ancients seemed to understand that memory and perception are intertwined—and they exploited this relationship with remarkable finesse. The asymmetrical wings of the Propylaea, for instance, achieve a precise visual balance only at the moment of arrival, when the structure opens itself to the spectator. The Caryatids are withheld until the proper moment, just as the Parthenon is revealed obliquely, rather than frontally—delaying its full impact in favor of a more nuanced spatial drama.

Eisenstein, in his interjections, likens these spatial effects to the calculation of a film shot: the power of the first visual encounter, yes, but also the deeper logic of sequential juxtaposition. To defend this argument, he sketches the compositional schemes for the four picturesque “shots” drawn from Choisy’s perspectival storyboard. The first two—the symmetrical view of the Propylaea and the axial unveiling of Athena Promachos—establish a rhythm of formal equilibrium. The third frame introduces a new asymmetry: the Parthenon, glimpsed obliquely, echoes the sculptural motif of the statue now left behind. The fourth scene—the Erechtheion and the Porch of the Caryatids—balances the composition from the opposite side. Together, these four views create a kind of visual polyphony, unified through the recurrence of sculptural themes and balanced spatial counterpoints.

Importantly, Eisenstein also draws attention to the temporal experience of architecture—not only what is seen, but how long each image is held before it yields to the next. The duration between frames, so to speak, mirrors cinematic rhythm. For Eisenstein, the length of a montage sequence is proportionate to the magnitude of the architecture itself: the scale of the buildings dictates the tempo of movement, the rhythm of perceptual unfolding. Architecture, in this sense, becomes a time-based art, not unlike cinema—structured by pauses, pacing, and progressions.

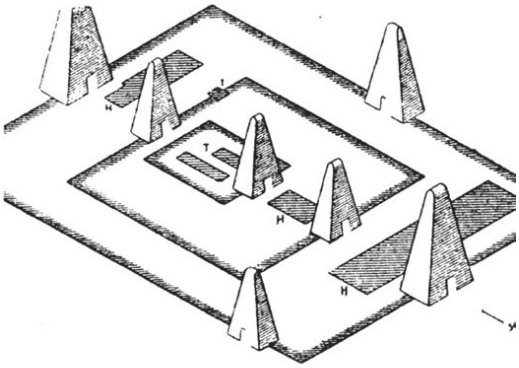
This choreographed engagement is not unique to the Athenian Acropolis. As Eisenstein suggests, such spatial montage is a universal architectural phenomenon. Across civilizations, monumental ensembles choreograph perception in similar ways—balancing near and far, inside and outside, concealment and revelation. Great works of architecture consistently activate the viewer’s body and mind in time, producing a cinematic experience long before cinema existed.



Eisenstein’s sketch : a-propylaea, b-athena promachos, c-parthenon, d-Erechtheion

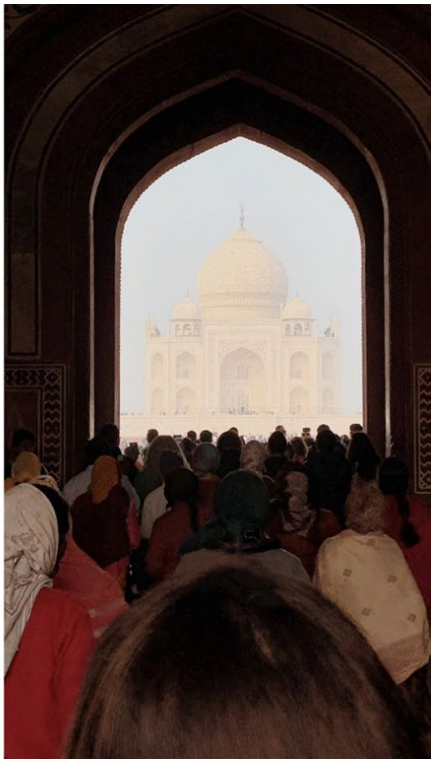
In revisiting the Acropolis through the lens of montage, Eisenstein uncovered a latent cinematic logic embedded in ancient architecture—a logic of guided perception, spatial rhythm, and sequential revelation. Drawing from Choisy's perspectival storyboard, he revealed how movement through space can be orchestrated like a film: the spectator's path unfolds like a montage, where architecture does not simply exist but performs.
Our return to the Acropolis, via digital simulation, extends this cinematic reading into a new medium. While simulation lacks the physical tangibility of architecture, it reactivates the sequence of spatial impressions with new potential. Virtual reconstructions, though speculative, become tools of spatial pedagogy—not because they represent "truth," but because they allow us to re-engage with architecture as an experience of movement, memory, and transformation.

Indian temple architecture offers a parallel spatial logic—one that unfolds through a choreographed procession and a carefully constructed sequence of thresholds. In the temple towns of South India, the towering gopuram functions not merely as a gateway but as an orienting landmark in the urban fabric. From a distance, it anchors the pilgrim's journey through the city, much like the glistening spear tip of Athena Promachos once guided sailors toward Athenian shores. Yet, as the visitor draws closer, the monument's perceptual totality dissolves. Like Athena's statue, the gopuram is no longer a beacon but a liminal threshold—announcing entry into a sacred domain.



Gopuram, Meenakshi Sundareswarar Temple

The processional approach to the Taj Mahal operates in a different cinematic fashion. As one navigates the dense, labyrinthine urban fabric of Agra, the monument remains hidden—its grandeur withheld until the final moment. Then, framed within the entrance archway, the Taj is revealed in full, staged as a singular, frontal elevation. This architectural sequencing—compression, anticipation, revelation—recalls Eisenstein's montage logic: the emotional and perceptual rhythm constructed through spatial delay and sudden clarity.

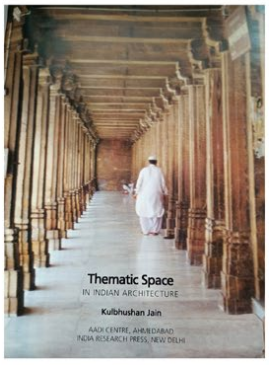


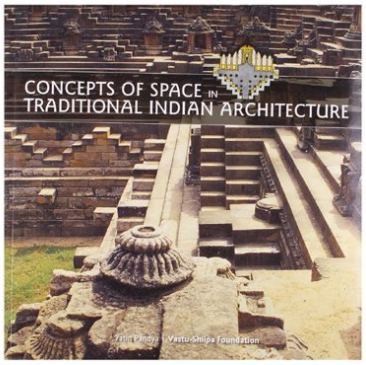
Taj Mahal

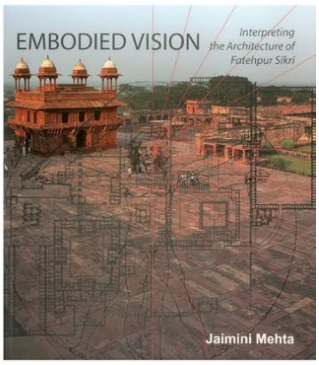
These spatial narratives resonate with a lineage of Indian architectural thought, notably Yatin Pandya's Concepts of Space in Traditional Indian Architecture, which reflects upon these transitions through typology and ritual use. Pandya's work can be seen as a continuation of Kulbhushan Jain's Thematic Space in Indian Architecture, and his earlier text, Elements of Space Making, which begins at the most elemental level—floor, wall, column, roof, aperture, stair. These foundational components coalesce into spatial compositions, which in Jain's work evolve into typological themes. Jaimini Mehta's Embodied Vision, in turn, extends these discussions to an urban scale, offering a phenomenological reading of Fatehpur Sikri that emphasizes the bodily and temporal dimensions of architectural experience.

Though these writings differ in scope and sequence, they can be read as a montage of architectural theory—each contributing a distinct yet interconnected shot in an intellectual sequence. Just as in cinema, meaning here emerges not in isolation but through juxtaposition, rhythm, and resonance.

This essay, then, does not aim to establish a new architectural theory, but to offer a lens of interpretation—reading Eisenstein's montage not only as a cinematic strategy, but as a mode of spatial understanding. Through the Acropolis and beyond—through digital simulation, temple towns, and texts on Indian space—montage reveals itself not as metaphor, but as method: a way of navigating, sequencing, and ultimately thinking architecture.







Books as sequential shots in indian architecture theory : Elements of space making(Yatin Pandya), Thematic space in Indian architecture (Kulbhushan Jain), Concepts of space in traditonal Indian architecture (Yatin Pandya) and Embodied vision interpreting the architecture of Fatehpur Sikri (Jaimini Mehta)

As the culmination of these mental notes and speculative interpretations, this compilation is intended as a reference point—a modest contribution toward a broader understanding of the perceptual qualities of architecture and the built environment. Many of the reflections gathered here are undoubtedly imperfect, even inaccurate, but they are offered not as definitive claims, but as part of an ongoing process of learning through interpretation.

The act of assembling these fragments was, above all, an attempt to gather dispersed thoughts into a coherent thread of inquiry. It is hoped that, over time, these notes will evolve—refined through further study, deepened by experience, and corrected where necessary. This is not a conclusive theory, but a scaffold for thinking, a framework open to revision.

As a closing gesture, I hope this compilation is received not as a proposal for a new architectural theory, but as a way of reading existing ones—transforming passive engagement into an active tool for design, critique, and deeper understanding. While the text may at times appear repetitive or digressive, its original intention was to render Sergei Eisenstein's essay more accessible by situating its ideas within a wider architectural context. Inevitably, much of Eisenstein's focus on iconography and symbolism remains outside the scope of this effort, which instead concentrates on his theoretical insight into spatial montage.

Ultimately, this is not a definitive interpretation, but an invitation—to read, to reflect, and perhaps, to reimagine architecture as a medium of cinematic and embodied experience.

**The author used LLM to refine the structure and grammar of this text.*

सांख्य

samkhya-in