Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Computational spatiality: virtual production, generative media and the politics of screen space

Rui Gao
(Zhejiang) Hengdian College of Film and Television
Yiwei Zhao
Longwood University
Weice Yang
156656000@qq.com

(Zhejiang) Hengdian College of Film and Television

Article

Keywords: computational spatiality, virtual production, generative media, screen space, visual culture, platform governance

Posted Date: September 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10478087/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Virtual production and generative media are often placed in different technical histories. One grows from real-time engines, tracking and computer graphics; the other from machine learning and statistical generation. On a screen, however, they now meet. Both change the conditions under which a location can be built, altered, approved and carried from one stage of production to another. This article develops computational spatiality as a concept for studying that change. The argument draws on film-space theory, visual culture, software and infrastructure studies, recent research on virtual production, and selected technical and regulatory documents. It is anchored in three visible configurations: virtual scouting for The Lion King, camera–environment reciprocity in StageCraft production, and spatial instability in video-generation systems. Virtual production usually works through persistent geometry. A tracked camera calls up a new view of a world whose coordinates and assets remain in place. Generative video works differently. It can deliver a convincing room, street or landscape without preserving every spatial relation behind the image. These are not rival stages in one technological sequence. They are distinct ways of making a screen world hold together. In both, creative authority moves beyond the composition of a frame. It also lies in setting parameters, controlling assets, choosing outputs, repairing continuity and deciding which version becomes public. Computational spatiality names this wider field. It helps explain why new freedom at the level of the image can arrive with stronger dependence on platforms, models, metadata, specialist labour and proprietary infrastructures.

1 Introduction

A camera moves towards an actor inside an LED volume. The horizon moves too. It changes its parallax for the tracked lens, and the change happens quickly enough to look natural. Light from the display falls across the actor’s face and catches the edge of a practical prop. Someone asks for dusk instead of afternoon. The request passes through a digital sky, a colour decision, a render budget and the limits of the wall. The place is present, but not in the old sense. It exists as one state that the system can run.

Now consider a generated street. It has wet pavement, shop windows and the shallow depth of field of a long lens. The camera appears to drift past the buildings. Yet the street may have no persistent geometry behind it. A doorway can slide between frames. A parked car may acquire another wheel. The road may bend where it did not bend before. The clip still offers the signs of a place. Whether it preserves the place is another question.

Film theory has never treated screen space as a transparent copy of the world. Framing selects. Mise-en-scène arranges bodies and things. Camera movement reveals some relations and hides others. Montage joins spaces that may never have shared a location (Heath, 1976; Bordwell, 1985). Digital visual effects made that construction more elaborate, but the finished image remained the main site at which heterogeneous elements were made to appear together (Prince, 2012). Virtual production and generative media shift part of the problem elsewhere. Before there is a finished image, a spatial state has already been built, called, revised, rejected or approved.

Several established concepts illuminate this shift. Digital space points to encoding, programmability and variability. Virtual space describes simulated environments and their relation to physical actuality. Immersive and navigable space place the body, orientation and movement at the centre of analysis. Procedural space explains variation produced through rules. Operational-image theory asks what images do inside systems, including operations that viewers never see (Manovich, 2001; Kitchin and Dodge, 2011; Parikka, 2023). Each concept matters. None quite captures the way a screen environment can persist across virtual scouting, asset construction, photography, post-production, delivery and later reuse while remaining open to controlled change.

This article calls that condition computational spatiality. The term refers to mediated space organised as a set of executable, revisable and governable possibilities. It does not describe every computer-generated image. A digital matte painting may be made with software and still function as a finished surface. The stronger condition appears when spatial relations can be called into action by a tracked camera, an engine, a simulation, a prompt, a model or a user input. The environment becomes more than an image file. It becomes a working field in which different states can be produced and compared.

The distinction matters because two different computational logics now shape screen production. Real-time virtual production usually depends on persistent geometry. Assets, coordinates, lenses and tracking systems maintain a world that can be rendered from changing viewpoints. Generative media can make persuasive spatial appearance through inference without guaranteeing the same persistence. The first gives geometry a live response. The second gives probability a visible form. They increasingly overlap in practice, but they do not make the same claim about the world behind the frame.

Three questions guide the discussion. What separates a computationally produced image from a computationally organised space? How do virtual production and generative media secure continuity in different ways? And who gains authority when the field of possible images is partly defined by engines, asset libraries, datasets, models, standards and platforms?

The answer developed here is that visual power now reaches further upstream. It lies in composing and framing, as it always has, but it also lies in deciding what can vary, which version can be run, what remains available for reuse and which output receives approval. Directors, cinematographers and production designers still make choices. So do technical artists, stage operators, software firms, model providers and rights holders. Their powers are not equal. A filmmaker may ask for a different valley. Another actor may decide whether such a valley exists in the asset library, whether the prompt is permitted, whether the model can sustain it and whether the result can be exported.

This is why computational spatiality belongs to more than film aesthetics. It joins questions from visual culture, production studies, software studies and the politics of infrastructure. It also describes a contradiction. The image becomes easier to alter in one place and harder to control in another. A sky can change in seconds. The pipeline that makes the change possible may have taken months to build.

2 Analytical approach and materials

The article is conceptual. It does not measure the prevalence of virtual production, estimate the effects of generative media or claim that one workflow represents the industry as a whole. Its task is narrower: to build a concept that can separate related phenomena without flattening their differences. Conceptual work is useful when a new object sits between established fields and when existing terms describe only parts of it (Jaakkola, 2020).

The argument was developed by moving between five bodies of material. The first includes work on cinematic space, narration and the social production of space. The second comes from software, interface and operational-image theory. The third concerns infrastructures and platforms. The fourth consists of peer-reviewed research on virtual-production practice, especially work that treats workflow and labour as more than technical afterthoughts. The fifth includes technical, corporate and regulatory documents that describe particular systems, capabilities or rules.

The selection was purposive rather than exhaustive. Sources were retained when they helped answer one of four practical questions. What holds the space together? What makes it change? What kind of labour keeps one version continuous with the next? Who can set the conditions under which the change occurs? These questions kept the reading close to the production process. They also prevented every use of software from being treated as evidence of a new spatial order.

The corporate sources need caution. Magnopus and Industrial Light & Magic provide unusually clear accounts of virtual scouting and StageCraft. OpenAI documents both the ambitions and the reported limits of its video systems. Such material is useful because it reveals how system builders describe their own architectures. It is not neutral evidence of creative benefit, cost reduction or democratisation. Promotional language tends to make friction disappear. The article therefore reads those claims beside peer-reviewed studies of workflow, coordination and labour.

Three configurations carry most of the analysis. The first is the virtual scouting system built for The Lion King. It shows how familiar cinematographic gestures can be translated into software operations. The second is the tracked relation between camera and LED environment in StageCraft. Here the environment responds to the apparatus that records it. The third is the spatial instability visible in generated video. This is not offered as a permanent catalogue of model failure. Systems improve. The instability matters because it makes the underlying spatial claim easier to see.

Breakdown is especially useful. A stage works best when the tracking chain disappears from attention. A generated shot works best when each frame seems to belong to one place. Tracking drift, stale assets, topology changes and lost metadata bring the hidden chain back into view. They show where continuity resides and who has to repair it. The article treats those moments as evidence about system structure rather than as comic mistakes.

The regulatory documents play a smaller role. Chinese rules on the labelling of generated content and the 2026 technical requirements for virtual-reality film show that provenance, content security and delivery can become formal parts of a visual object’s public life. They do not prove a universal regulatory

trend. They show one concrete setting in which origin and responsibility are being built into the circulation of images.

The framework is judged by three modest standards. It should distinguish computational spatiality from neighbouring concepts. It should account for both persistent geometry and probabilistic appearance without pretending they are the same process. It should also lead to empirical questions that later research can test through ethnography, contract analysis, infrastructure mapping or close examination of generated sequences. A concept is useful only if it gives research somewhere to go.

There is also a reason to keep the examples uneven. The Lion King system and StageCraft leave a relatively clear technical trail, while generated video is documented through model reports, public demonstrations and visible outputs. Treating the materials as if they were equivalent would create a false balance. The comparison rests instead on the relations each source can show. Production-company accounts can describe how a camera is tracked or how an environment is assembled. They cannot tell us, on their own, how widely a workflow has been adopted or how workers experience it. Model reports can identify stated capabilities and known limits. They cannot substitute for independent testing.

This asymmetry is part of the object, not an inconvenience to be hidden. Contemporary screen infrastructures are disclosed selectively. Some operations are public because they sell a product. Others remain inside contracts, proprietary code and production routines. A conceptual synthesis must therefore work with partial views while marking where the evidence stops. Throughout the article, technical capacity is kept separate from social use. A system may permit a form of control without exercising it in every production. It may also acquire uses that its developers did not foresee.

3 Defining computational spatiality

Screen space has always taken work. Heath's account of narrative space describes the labour by which discontinuous fragments become an intelligible field (1976). Bordwell also treats narration as patterned cueing rather than the delivery of a ready-made world (1985). Computation does not remove framing, performance, editing or off-screen space. It changes part of the backstage. A coherent location may now depend on engine scenes, asset branches, lens metadata, calibration files, model checkpoints, moderation rules and version histories. None needs to appear in the final shot.

Lefebvre's argument that space is socially produced keeps this backstage from looking neutral (1991). A studio set, a virtual city and a generated room all condense institutions, labour and cultural habits. Computation gives some of those relations an operational form. A location can be split into assets. A camera coordinate can trigger a new perspective. A prompt can call familiar architectural types into view. Space is still represented. It is also processed.

Software studies gives names to several parts of this process. Manovich identifies programmability, variability and database logic as central features of new media (2001, 2013). Kitchin and Dodge show

that code can help constitute a space rather than merely describe it (2011). Galloway's interface is not a transparent window but a relation that makes some actions possible and keeps others out of sight (2012). Parikka extends the discussion to operational images whose decisive work may remain invisible to a spectator (2023). Computational spatiality draws from each line of thought, but its object is larger than a single interface or image. It concerns the life of spatial relations across a production chain.

The concept begins with an operational substrate. Location, scale, perspective, lighting, material and movement must become adjustable in some form. They do not all need to be numerical, and they do not have to sit in one program. They do need to be available to action. A camera can call a view. A lighting state can be restored. An asset can be replaced without rebuilding the whole scene. A model can infer another version from a prompt and a reference image.

Time enters next. Some systems are live. They couple spatial change to a camera, performer or operator while the image is being made. Other systems are generative. They produce states that were not fully specified in advance. The distinction is not absolute. A generative asset can be loaded into a real-time engine, and a tracked stage can include conventionally rendered material. Still, liveness and generativity place uncertainty in different parts of the process. The stage has to keep a prepared world aligned. The model has to keep an inferred world from drifting apart.

Authority sits inside both. One person may imagine a change. Another makes it executable. Someone else selects the surviving version. A company may own the engine, the model or the asset repository on which all of them depend. Computational spatiality is therefore political from the start. Its politics does not begin only when a platform bans an image or when a regulator demands a label. It begins earlier, in the design of the field from which images can be produced.

A threshold follows. Using Blender, Maya or a digital matte painting does not automatically create computational spatiality in the strong sense. The threshold is crossed when spatial states become operative across relations: a camera coordinate changes the background; an environment persists as a shared asset; a model returns several versions from incomplete instructions; a version has to remain traceable through editing and delivery. The question is not whether a computer was used. It is whether space has become a state that can be run, altered and governed.

Figure 1 places this condition within a wider field. Represented screen space remains relatively close to fixed presentation, even when it is digitally produced. Navigable immersive space makes orientation and movement variable. Real-time virtual production combines persistent geometry with live execution. Generative media moves towards probabilistic appearance. These positions overlap. A single production may use all four, sometimes in the same shot.

Table 1 clarifies these conceptual boundaries by showing how computational spatiality differs from digital space, virtual space, immersive or navigable space, procedural space and the operational image.

Table 1
Conceptual boundaries of computational spatiality

Concept	Primary object	What it explains	Question left open
Digital space	Digitally encoded or processed space	Programmability, variability and database structure	How spatial states are executed, versioned and authorised across production
Virtual space	Simulated space in relation to physical actuality	Relations between actual and simulated environments	Hybrid events in which physical actors, tracked cameras and rendered environments co-produce an image
Immersive or navigable space	Embodied orientation and movement	Presence, attention and path-dependent experience	The production infrastructures that make spatial states available
Procedural space	Space generated through rules	Rule-based variation and emergence	Camera tracking, asset governance, trained models and cross-department versioning
Operational image	Images that perform tasks within systems	Action, calculation and operations that may remain invisible	The persistence of a coordinated environment across a production chain
Computational spatiality	A field of executable and governable spatial possibilities	Execution, versioning, generation and unevenly distributed spatial agency	Must be specified through situated systems, labour and power relations

4 Executable geometry in virtual production

Virtual production is often pictured through its hardware: the LED wall, the tracking marker, the motion-capture stage. The picture is incomplete. The deeper change is the conversion of place into a callable asset system. A mountain can become a mesh, a material set, a lighting profile, an atmospheric layer and a versioned object. Its distance from the camera can change. Its weather can change. The mountain has not become immaterial. It has become addressable.

The virtual-production system developed for The Lion King makes this clear. Magnopus describes a multiplayer environment in which filmmakers entered a computer-generated landscape through virtual reality and used tools modelled on familiar cinematographic equipment (Magnopus, n.d.). The phrase “inside the film” belongs to promotion. The actual operation is more interesting. The filmmakers entered an interface to terrain, light, animals and viewpoints. The virtual camera became a query. Move it, and the system returned another relation among those elements.

This did not simply restore live-action practice inside a computer. It translated practice. A director could still walk, look and frame. Yet the gesture now passed through a tracked controller, a database of assets

and software rules. The body remained important, but it was no longer dealing with an indifferent physical location. It was working through an environment built to respond.

StageCraft adds a stronger form of reciprocity. Industrial Light & Magic describes perspective-correct backgrounds rendered for a live-tracked camera and displayed on an LED volume (Industrial Light & Magic, 2020, n.d.). In ordinary location photography, the camera moves and the place does not recalculate itself. On a tracked stage, the visible environment changes for the camera. The lens does more than choose a view. Its position helps produce the view that exists at that moment.

The effect is persuasive when many small agreements hold. Camera position, lens data, tracking latency, colour, exposure, display refresh and render performance must line up (Chen and Zhao, 2023). The foreground and background must share a plausible light. Moiré has to stay under control. Reflections must not reveal the wrong angle. Liveness is not an escape from mediation. It is mediation kept close enough in time to disappear.

This has consequences for labour. Departments can work against the same evolving environment during photography. Production design, cinematography, visual effects and performance meet earlier and more often. That can reduce some hand-offs. It can also create new ones. A production designer’s request may need to be translated into geometry, texture, optimisation and render priority before it can be seen on the wall.

Recent production research warns against the fantasy of a universal workflow. Virtual production remains project-specific. It requires early feasibility work, technical rehearsal and negotiation among people who do not always share the same vocabulary (Swords and Willment, 2024; Dean et al., 2026). A change that looks trivial in a meeting may demand a rebuilt asset or a new lighting strategy. The stage is flexible after a great deal has already been fixed.

Failure exposes the arrangement. Tracking drift breaks the agreement between camera and horizon. A stale asset branch can return the wrong doorway or the wrong skyline. A reflection can reveal the edge of what the wall can provide. These are not merely ugly images. They are failures of a spatial claim. The shot asks the viewer to believe that the perspective belongs to a coherent world. When the chain slips, the world loses its footing.

Virtual production therefore redistributes certainty. More can be previewed. Some decisions move closer to the set. Yet preproduction grows heavier. Virtual art departments, asset management, colour pipelines and contingency plans move forward in time. A sky can be changed fast because someone prepared the sky, named it, tested it and made it available. The speed is real. So is the work hidden inside it.

Versioning gives this workflow a memory. A shot can return to an earlier engine state, compare two lighting setups or restore an approved asset branch. That memory is not automatic. Names, dependencies and permissions must be maintained. When they are not, the set may still render, but

nobody can be certain which set it is. The practical value of persistent geometry therefore lies as much in recoverability as in visual realism. A world can be revisited because its relations have been kept.

The arrangement also changes rehearsal. On a physical location, the crew learns what the place permits by moving through it. In a virtual environment, part of that learning happens before the stage is occupied. Lens choices, blocking and light can be tested against a model, but the test may also harden assumptions. An early asset becomes the place against which later decisions are made. Previsualisation can open the work. It can also close it sooner than anyone notices.

5 Probabilistic appearance in generative media

Generative video works from another premise. A conventional 3D environment normally retains coordinates and object identities when the camera turns away. A text-to-video model can produce the appearance of such a world without storing it in the same form. It draws on learned regularities, prompts, references and prior frames. The result can look photographed. Behind the look may be a sequence of local inferences.

OpenAI's 2024 report presented Sora as a step towards models that simulate aspects of the physical world. It also listed problems involving causality, object persistence, interaction and spatial consistency (OpenAI, 2024). Sora 2 was later described as more physically accurate and more controllable (OpenAI, 2025). That change is important. Model limits move. The argument here does not depend on any one defect surviving forever.

What remains is a difference in how persistence is secured. A realistic image does not by itself prove that the model maintains the same room behind every camera move. Spatial continuity has to be produced through inference, memory, conditioning or some other mechanism. It cannot be assumed simply because the floor has texture and the light falls in the right direction.

A prompt is a thin spatial instruction. It can name a city, a lens, a time of day and a camera movement. It cannot describe every distance, surface and relation. The model supplies the rest from patterns sedimented in image collections. MacKenzie and Munster's account of platform seeing is useful here because it places the single image inside larger computational ensembles (2019). A generated street is not created by language alone. It draws on learned conventions about streets, weather, wealth, danger and human presence.

The geography is probabilistic in this limited sense. It is assembled from likelihoods rather than guaranteed by a stable map or set. Cinema has always invented places. That is not new. What changes is where internal consistency comes from. In a built set or a modelled environment, many spatial relations exist before the shot. In generated video, some of those relations may need to be inferred again as the sequence moves.

This is why the glitch matters. A distorted hand is a familiar sign of generation. A corridor that changes width tells us more about space. It shows that the clip can preserve the idea of a corridor while losing part of the corridor's geometry. The viewer does repair work. We read successive frames as one environment until the differences become too large.

The glitch is not sacred, and it is not proof that models cannot learn better world consistency. It is evidence of the contract under strain. When the contract holds, visual plausibility stands in for a stable world. When it breaks, we see the gap. The image had depth. The place did not fully persist.

Continuity work changes accordingly. A generative system may preserve colour and texture while losing scale. It may keep a face and alter the room. It may preserve a camera style while changing the object that the camera follows. Artists respond with reference images, masks, repeated prompting, selection, compositing and conventional post-production. The workflow is less like pressing a button than the rhetoric of automation suggests.

This labour is selective and often invisible. Many outputs are discarded. Some are repaired frame by frame. Others become reference material for another pass. Dataset builders, model engineers and moderation systems have already shaped the space of possibility before the filmmaker begins. Crawford and Paglen show how training sets carry classificatory politics, while Crawford places AI within material and institutional infrastructures (Crawford and Paglen, 2019; Crawford, 2021). The generated city arrives with a history, even when that history cannot be inspected.

The boundary of a stage is easier to see than the boundary of a learned distribution. The LED wall ends. The model's field of likelihood does not announce where it thins out. Some places appear readily. Others need careful references, local data or repeated correction. Part of spatial authorship has moved into a system whose limits become visible through outputs, refusals and recurring mistakes.

Table 2 summarizes the contrast between executable geometry in virtual production and probabilistic appearance in generative media across substrate, viewpoint, truth claim, failure mode, continuity labour, dominant gatekeepers and political tension.

Table 2
Two computational logics of screen space

Analytical dimension	Executable geometry: virtual production	Probabilistic appearance: generative media
Spatial substrate	Persistent 3D assets, coordinates and calibrated engine state	Learned distribution, input conditions, references and repeated inference
Camera or viewpoint	Tracked input recalculates perspective within a coordinated world	Prompted or inferred movement may imply depth without a persistent world model
Primary truth claim	Visible perspective corresponds to a coherent coordinate chain	Local visual plausibility can be sufficient to imply a world
Characteristic failure	Tracking, calibration, render or asset-version mismatch	Object mutation, topology drift, causal or geometric inconsistency
Continuity labour	Synchronising assets, engine states, lens data and approved takes	Constraining, selecting, repairing and tracing repeated generations
Dominant gatekeepers	Engine ecosystems, stage access, asset owners and technical specialists	Model providers, training distributions, interfaces, moderation and compute access
Political tension	On-set plasticity alongside infrastructural lock-in	Apparent abundance alongside opaque limits and cultural defaults

Time is a harder problem than a single frame suggests. A generated room may look stable for two seconds and fail in the fifth. A short clip can hide drift that becomes obvious when a scene needs coverage, reverse angles or later revisions. Production continuity therefore depends on more than the quality of an isolated output. It depends on whether a place can survive duration, return and comparison. The problem grows when the same environment must support several shots rather than one striking image.

Narrative space adds another test. A doorway can move and still pass unnoticed if no character uses it. Once a character crosses the room, spatial relations become part of the action. The environment must remember where objects are and what they allow. This is one reason generated video often appears strongest in brief, self-contained moments. A longer scene asks the image to carry obligations forward.

6 Governing spatial possibility

The politics of computational spatiality begins with a simple fact: more choices do not mean equal control over the field of choice. A filmmaker may receive hundreds of variants and still have little say over the model that produced them. A stage may permit rapid changes while keeping its assets inside a proprietary environment. Governance here means the allocation of the power to vary, execute, select, approve and trace a spatial state.

Those powers appear at different moments. Conditioning comes first. Parameters, assets, prompts, training data and safety rules establish what may vary. Execution turns those conditions into a tracked render or a generated output. Selection keeps some versions and discards others. Ratification fixes one version for editing, delivery or exhibition. It may also attach rights information and provenance. In a small production, the same person can perform several of these acts. In a larger one, they are split across departments and firms.

The stages of this governance process are mapped in Fig. 2, which shows how spatial possibility is conditioned, executed, selected and ratified across conceptual, operational, selective and infrastructural forms of authority.

The split can be mundane. A director asks for a valley to feel narrower. In a real-time environment, the request may become a change in camera position, focal length, terrain, atmosphere or light. In a generative workflow, it may become a rewritten prompt, a reference image, a mask and another batch. The request survives, but it changes form as it moves. By the time the valley appears, the system has translated an aesthetic judgment into available operations.

Interfaces do not merely speed up that translation. They shape it. Platform and infrastructure studies show how standards, classifications and access conditions organise participation before any visible prohibition occurs (Nieborg and Poell, 2018; Plantin et al., 2018). Engines decide which material systems are supported. Asset stores decide what is searchable. Model interfaces decide which controls a user can see. Cloud services decide how files move and how long histories remain available.

This creates a politics of default. A shared engine can make certain kinds of vegetation, weather and architecture easy to deploy. A training distribution can make some interiors or urban forms statistically familiar. The result need not be visual sameness. Defaults work through effort. Some spaces arrive quickly. Others cost more time, data and expertise.

A generic city street may be produced in seconds. A locally specific street can be harder. Its paving, shop signs, scale and relation between public and private space may not fit the patterns that the model or asset library offers readily. The difficulty is not simply technical. It reflects what has been photographed, labelled, modelled, licensed and maintained.

Labour reorganises around these conditions. Virtual art departments build environments before the shoot. Technical artists keep geometry, lighting and hardware in balance. Stage operators translate between the camera crew and the engine. Generative workflows add prompting, ranking, repair and provenance work. None of this is a minor supplement. It is the labour that keeps possibility usable.

Authorship becomes conditional rather than absent. The filmmaker can still set direction, refuse an output and rebuild a scene. But the range of workable choices is shared with systems designed elsewhere. Bucher’s account of algorithmic power is helpful because power often operates by making

some actions easier, more likely or more intelligible than others (2018). The author composes within a field whose edges may belong to another organisation.

Provenance makes the field visible in another way. China's 2025 rules on labelling AI-generated and synthesised content require visible and metadata-based forms of identification for generated media (Cyberspace Administration of China et al., 2025). The 2026 technical requirements for virtual-reality film also connect content, exhibition, security and delivery (National Film Administration of China, 2026). These measures belong to a specific regulatory setting. They should not be treated as a universal model.

Even so, they show that an image may need more than a visible surface in order to circulate. It may need a record of origin, modification and responsibility. Provenance then becomes part of the image's operational existence. It can support accountability. It can also create new forms of surveillance and gatekeeping. The same metadata that proves where an asset came from may reveal who altered it, when and under whose authority.

Rights travel unevenly through the same field. A background plate, a 3D asset, a captured performance and a generated voice may sit in one shot while remaining subject to different contracts. Technical composability can make them look like one object. Legally and institutionally, they are not. A production may be able to call an asset again and still lack the right to reuse the performance inside it. Governance becomes visible when the file works but permission does not.

Small productions face another asymmetry. They may gain access to tools once reserved for large studios, yet they often enter on rented terms. Storage, model calls, asset licences and export options can change with the service. The barrier to making an image falls. The barrier to keeping control of the image may remain. This is not an argument against shared infrastructure. It is a reason to study its terms closely.

7 Infrastructural continuity and visual truth

Classical continuity is usually discussed as a relation among shots: screen direction, gesture, lighting, position and narrative time. Computational production adds another layer. Assets, engine states, lens records, model outputs, approvals and delivery formats also have to remain aligned. A visible error may begin far from the edit.

A wrong environment branch can place an actor in yesterday's version of a room. A stale calibration file can make a wall seem to slide. A generated background can change the location of a window. Provenance can disappear when a file moves between systems. The final image carries the result, but the cause may sit earlier in the chain.

Infrastructural continuity names the work that keeps the chain intact. It does not replace aesthetic continuity. It makes aesthetic continuity reproducible. In virtual production, the claim rests on the relation among assets, engine state, camera and approved take. In generative media, references, constraints,

selection and repair do more of the work. Both require a spatial state to remain identifiable long enough to be accepted.

As shown in Fig. 3, infrastructural continuity depends on the alignment of spatial substrate, execution state, view condition, approved version and delivery state before the image reaches the public.

The two systems therefore support different regimes of visual truth. Virtual production asks the spectator to accept that the photographed perspective corresponds to a coordinated environment, even if that environment is rendered. Its truth lies in the calibrated relation among lens, geometry, display light and performance. Generated video can ask for less. Local plausibility may be enough to imply a world.

This complicates older debates about indexicality. In an LED volume, physical light reaches the sensor, but part of that light comes from a responsive image. A performer occupies a real set while the horizon is calculated for the camera. In generated video, no photographed event may exist at all, yet the clip can reproduce the grammar of lenses, depth of field and camera movement. The question is no longer only where the image came from. It is whether the spatial relations it claims can be trusted and traced.

The spectator rarely sees the labour that answers that question. We see a stable room, not the asset identifiers and version histories behind it. We see a generated street, not the rejected clips that failed to keep the street intact. Crary's history of the observer shows that visual subjects are formed through apparatuses and practices rather than standing outside them (1990). Computational spatiality extends this insight into production. The spectator meets the achieved result of a distributed system.

A glitch can break immersion, but it can also expose the system's promise. Broken parallax reveals a failed relation between camera and world. A changing object reveals a failure to preserve identity across inference. These errors need not be celebrated. They are useful because they tell us what kind of continuity the apparatus was trying to secure.

The audience can sense these differences without knowing their source. A tracked background that holds perspective gives the shot a quiet solidity. A generated space that drifts can feel dreamlike, comic or wrong. Such responses are not reliable technical diagnostics, but they matter. Visual truth is lived as confidence before it is explained as a pipeline. The infrastructure succeeds when the viewer no longer needs to ask whether the room will remain a room.

8 Discussion

Computational spatiality contributes by joining three questions that are often separated. How is space arranged inside an image? How does that space persist as production moves forward? Who controls the possible states from which the image is chosen? Film theory has long answered the first question. Software and production studies help with the second. Infrastructure and platform studies make the third harder to ignore.

The concept does not replace digital, virtual, immersive, procedural or operational space. It cuts across them. Its object is the environment that can be run, revised and governed. This is why a finished computer-generated landscape is not enough. The landscape becomes computationally spatial when camera data, asset states, prompts or model conditions can call different versions into being and when those versions have to survive a chain of decisions.

The comparison between persistent geometry and probabilistic appearance gives the concept its edge. Virtual production makes a prepared world responsive. Generative media makes an inferred world visible. One seeks continuity through coordinates and calibrated state. The other may achieve continuity through learned patterns, references and repeated inference. Hybrid workflows will blur the boundary, but they will not erase the practical decision beneath it: which spatial relations need to be secured and which may remain plausible approximations?

That decision carries cost. Geometry takes time to build and optimise, yet it can provide repeatable viewpoints and stable object relations. Generation can produce abundance quickly, yet abundance shifts work towards selection, repair and control. Neither route is simply more free. Each moves effort to another place.

The contradiction is clearest in the language of instant change. On a virtual stage, a sky can be altered during the shoot. The speed is impressive. It rests on prepared assets, trained operators, reliable tracking and enough computing power to keep the result live. In a generative system, a city can appear from a sentence. The city rests on data, model design, moderation and infrastructure that the user may never inspect. Local plasticity grows. Structural sovereignty may shrink.

This does not mean that platforms or vendors determine every image. Creative workers bargain with systems. They make local copies, build custom assets, combine tools and keep manual checks. They also refuse. A cinematographer can reject a wall solution. A production designer can insist on a specific place. An editor can return to conventional compositing. Agency remains, but it now includes the work of maintaining an exit.

Exit matters because assets and histories have weight. A production can leave one tool and still lose the version records, permissions or metadata needed to continue. A generated character can move as an image while its voice rights, reference files or moderation status stay elsewhere. Portability is therefore more than file export. It is the ability to carry the working memory of a spatial state into another system.

The cultural stakes are also uneven. Computational systems make some environments cheap to imagine because those environments are already dense in datasets and asset stores. Others remain thin. A model may know the visual shorthand of a generic metropolis and struggle with a village whose buildings, street life and material textures are poorly represented. An asset library may contain many medieval European walls and few regionally specific alternatives. The gap appears as extra labour.

This matters for visual culture because repetition does not need to look identical in order to be political. A default can survive through small choices: the available façade, the familiar skyline, the skin material that renders well, the weather system that already works. Over time, technical convenience can acquire the force of common sense. What is easiest to produce begins to look natural.

There is no simple policy fix. Open standards can reduce lock-in, but standards also decide what information counts. Provenance can support trust, yet detailed logs can become tools for worker surveillance. Asset reuse can lower costs, yet rights may be detached from the people whose performances or likenesses made the asset valuable. Governance has to distinguish several kinds of data that are often bundled together: data needed to keep a shot technically continuous, data used to evaluate labour, data used to train a model and data disclosed to the public.

For production research, the framework suggests a concrete agenda. Ethnography can follow a spatial state as it moves from design to set and post-production. Contract analysis can trace who owns digital environments, performer replicas and derived metadata. Infrastructure mapping can reveal where files, permissions and model calls pass between firms. Technical studies can measure spatial drift without reducing it to a single quality score.

The framework also invites regional comparison. Asset libraries and training distributions are not geographically neutral. Labour institutions differ. So do copyright rules, standards and public investment in production infrastructure. The same tool can support a proprietary platform in one setting and a shared public facility in another. Computational spatiality should therefore be studied as a situated arrangement, not a universal stage of media development.

The limits of the present article are clear. The examples are visible and well documented, which also makes them unusual. Corporate accounts foreground successful systems. They say less about ordinary bargaining, failed shoots and abandoned workflows. The article concentrates on visual space. Sound, haptics and the bodily experience of performers deserve fuller treatment. Video-generation systems also change quickly, so a failure described today may be reduced tomorrow.

Those limits do not remove the central problem. They show where the concept must meet evidence. Future work should ask when a generated environment becomes stable enough to reuse, how crews decide which relations need geometry, who pays for repair and what is lost when a project crosses from one infrastructure to another. These are practical questions. They are also questions about power.

9 Conclusion

Virtual production and generative media make screen worlds in different ways. Virtual production gives a modelled environment a live response to cameras, performers and engines. Generative media produces spatial appearance through inference and may imply a world without preserving every relation that would make the world stable. Both move important decisions away from the finished frame.

Computational spatiality names the field in which those decisions occur. Space is parameterised, executed, generated, revised and approved across infrastructures. Creative possibility can grow inside that field. Dependence can grow with it.

The image is no longer only a surface to be read. It is also the public face of a chain. The chain includes assets, coordinates, prompts, models, permissions, workers and records. To study contemporary screen space, we must ask what kind of world the image claims, how that world is kept together and who is allowed to decide what survives.

Declarations

Ethics Statement

This study is a conceptual analysis based on published literature, corporate and technical documents, and regulatory materials. It did not involve human participants, human data, animal subjects or personal data. Ethical approval and informed consent were therefore not required.

Author Contribution

Author ContributionsConceptualization: Rui GaoMethodology: Rui Gao and Yiwei ZhaoInvestigation: Rui GaoData curation: Rui GaoFormal analysis: Rui Gao and Yiwei ZhaoVisualization: Rui GaoFunding acquisition: Rui Gao and Weice YangSupervision: Yiwei ZhaoWriting – original draft: Rui GaoWriting – review & editing: Yiwei Zhao, and Weice Yang

Data Availability

This study is a conceptual and interpretive synthesis and did not generate or use an original quantitative dataset, transformed dataset, statistical variables, formal coding dataset, statistical formulae, or computational or statistical analysis scripts. Supporting documentation underlying the analysis has been deposited in Zenodo at https://doi.org/10.5281/zenodo.21892142. The deposited materials include: (1) a source–argument traceability matrix documenting the scholarly, technical, corporate, developer, and regulatory sources formally used in the manuscript and the specific propositions they support; (2) a register of selected contextual materials from a commercial industry report consulted during research development, including relevant page ranges, the source basis stated in the report, and the relationship of those materials to the manuscript; and (3) an analytical provenance map documenting the analytical questions and comparative reasoning through which the principal conceptual distinctions were developed. The commercial report's numerical estimates were not treated as independently verified empirical findings in the manuscript. Copyright-protected books,commercial industry report subscription-based journal articles, and other third-party materials for which redistribution rights are not held are not reproduced in the repository; complete citations, DOIs, stable

URLs, and access information are provided where available to enable independent scrutiny of the sources.

References

1. Bordwell D (1985) Narration in the fiction film. University of Wisconsin, Madison, WI
2. Bucher T (2018) If… then: algorithmic power and politics. Oxford University Press, Oxford
3. Chen J, Zhao J (2023) Film virtualisation production [Dianying xunihua zhizuo]. Tsinghua University, Beijing
4. Crary J (1990) Techniques of the observer: on vision and modernity in the nineteenth century. MIT Press, Cambridge, MA
5. Crawford K (2021) Atlas of AI: power, politics, and the planetary costs of artificial intelligence. Yale University Press, New Haven, CT
6. Crawford K, Paglen T (2019) Excavating AI: the politics of images in machine learning training sets. AI Now Institute. https://excavating.ai/. Accessed 25 July 2026
7. Cyberspace Administration of China, Ministry of Industry and Information Technology, Ministry of Public Security, National Radio and Television Administration (2025) Measures for labelling artificial intelligence-generated and synthesised content. https://www.cac.gov.cn/2025-03/14/c_1743654684782215.htm. Accessed 25 July 2026
8. Dean L, Gilardi F, Jamil S, Tu S, Zhu Z (2026) Developing an industrial workflow for virtual production filmmaking. https://doi.org/10.1177/13548565261433131. Convergence
9. Galloway AR (2012) The interface effect. Polity, Cambridge
10. Heath S (1976) Narrative space. Screen 17(3):68–112. https://doi.org/10.1093/screen/17.3.68
11. Industrial L, Magic (2020) Groundbreaking LED stage production technology created for hit Lucasfilm series The Mandalorian. https://www.ilm.com/groundbreaking-led-stage-production-technology-created-for-hit-lucasfilm-series-the-mandalorian/. Accessed 25 July 2026
12. Industrial, Light, Magic (n.d.), StageCraft (eds) https://www.ilm.com/stagecraft/. Accessed 25 July 2026
13. Jaakkola E (2020) Designing conceptual articles: four approaches. AMS Rev 10:18–26. https://doi.org/10.1007/s13162-020-00161-0
14. Kitchin R, Dodge M (2011) Code/space: software and everyday life. MIT Press, Cambridge, MA
15. Lefebvre H (1991) The production of space (trans: Nicholson-Smith D). Blackwell, Oxford
16. MacKenzie A, Munster A (2019) Platform seeing: image ensembles and their invisualities. Theory Cult Soc 36(5):3–22. https://doi.org/10.1177/0263276419847508
17. Magnopus (n.d) The Lion King: virtual production. https://www.magnopus.com/projects/the-lion-king. Accessed 25 July 2026
18. Manovich L (2001) The language of new media. MIT Press, Cambridge, MA

19. Manovich L (2013) Software takes command. Bloomsbury Academic, New York
20. National Film Administration of China (2026) Virtual reality film, Part 1: general technical requirements (DY/Z 14.1–2026). https://www.chinafilm.gov.cn/xxgk/gztz/202602/t20260209_949570.html. Accessed 25 July 2026
21. Nieborg DB, Poell T (2018) The platformization of cultural production: theorizing the contingent cultural commodity. New Media Soc 20(11):4275–4292. https://doi.org/10.1177/1461444818769694
22. OpenAI (2024) Video generation models as world simulators. https://openai.com/index/video-generation-models-as-world-simulators/. Accessed 25 July 2026
23. OpenAI (2025) Sora 2 is here. https://openai.com/index/sora-2/. Accessed 25 July 2026
24. Parikka J (2023) Operational images: from the visual to the invisual. University of Minnesota Press, Minneapolis, MN
25. Plantin J-C, Lagoze C, Edwards PN, Sandvig C (2018) Infrastructure studies meet platform studies in the age of Google and Facebook. New Media Soc 20(1):293–310. https://doi.org/10.1177/1461444816661553
26. Prince S (2012) Digital visual effects in cinema: the seduction of reality. Rutgers University Press, New Brunswick, NJ
27. Swords J, Willment N (2024) The emergence of virtual production: a research agenda. Convergence 30(5):1557–1574. https://doi.org/10.1177/13548565241253903

Figures

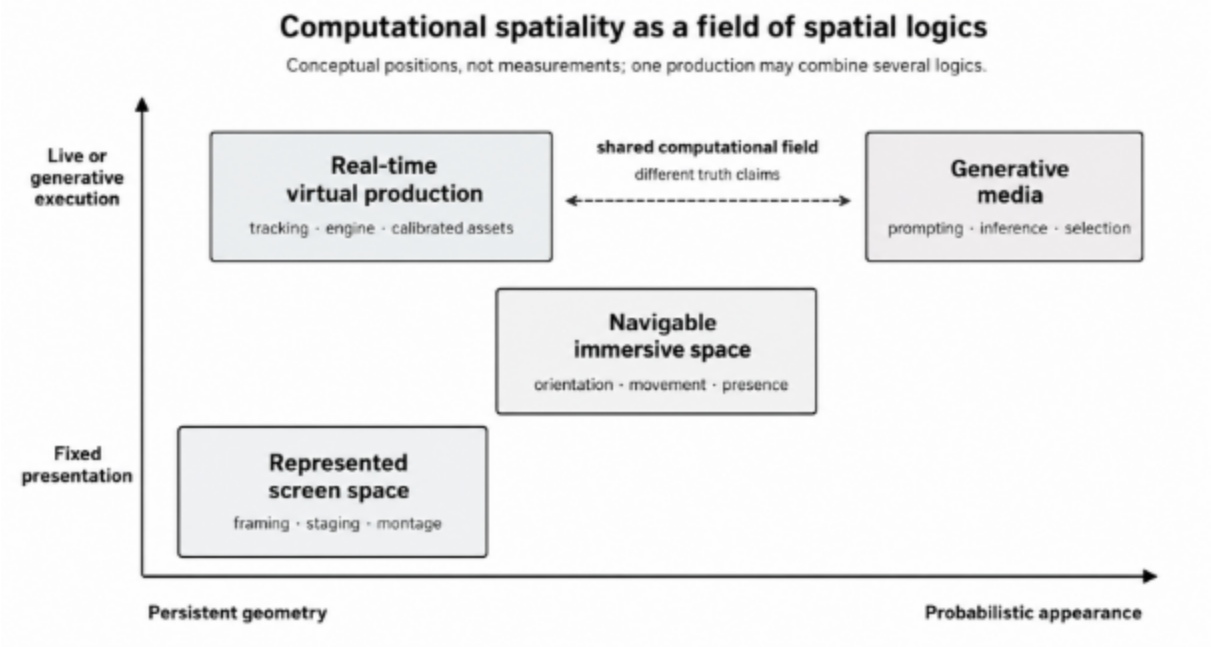


Figure 1

Computational spatiality as a field of spatial logics. The diagram contrasts persistent geometry with probabilistic appearance and fixed presentation with live or generative execution. The positions are conceptual, and one production may combine several logics.

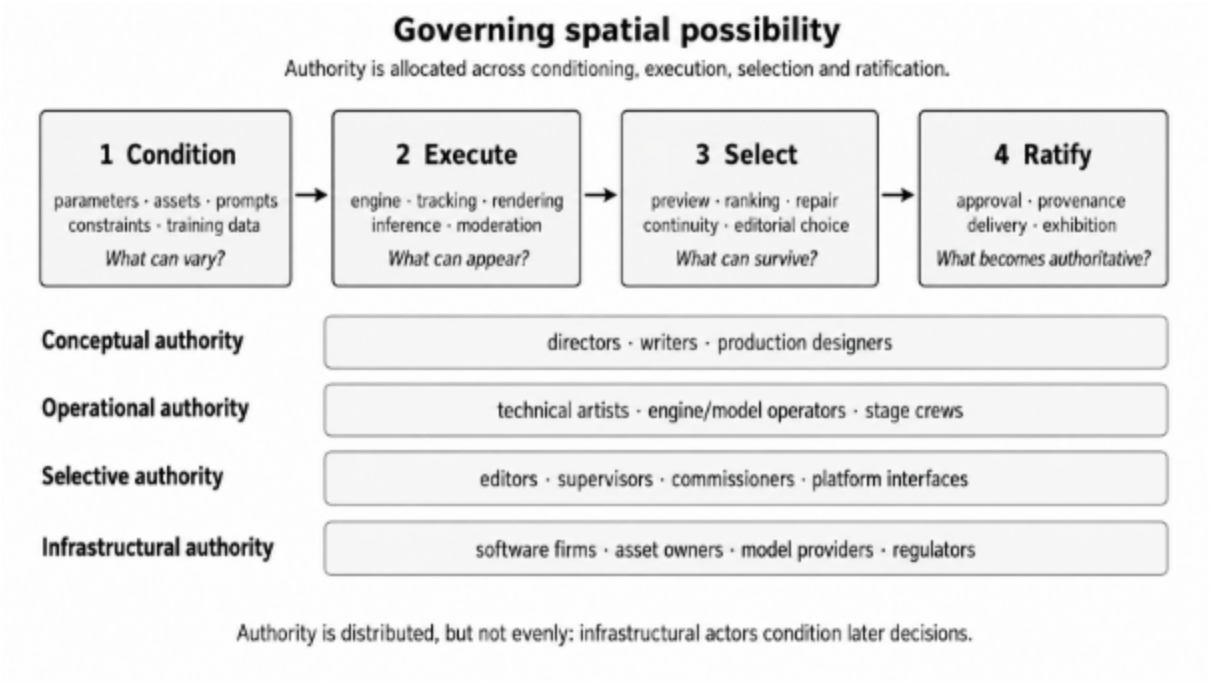


Figure 2

Governing spatial possibility. Authority is distributed across conditioning, execution, selection and ratification. Infrastructural actors influence the range of available options at every stage.

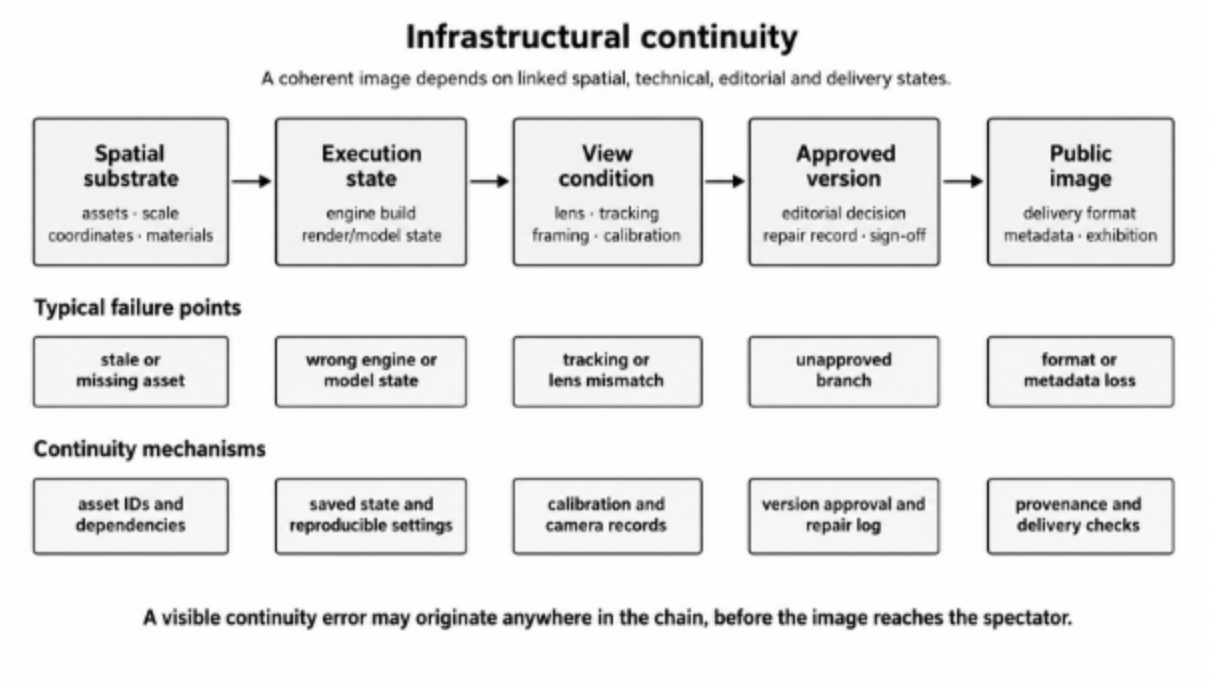


Figure 3

Infrastructural continuity. A coherent public image depends on continuity across spatial substrate, execution state, view condition, approved version and delivery. A visible error may begin anywhere in the chain.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupportingMaterialsfortheConceptualAnalysisofComputationalSpatiality.xlsx
- SupplementaryInformation.docx