Geodesics of Attention

Cinema, Gesture, and the Geometry of Perception

Flyxion

March 13, 2026

Abstract

We propose a unified geometric framework in which cinema, painting, drawing, swipe-gesture input, and visual perception are instances of a single underlying structure: directed motion through a structured spatial or perceptual field. A film is modeled as a continuous path through a scenario manifold $\mathcal{S} = \mathcal{C} \times \mathcal{A} \times \mathcal{L} \times \mathcal{E}$ whose factors describe camera configuration, actor pose, lighting, and environment. The perceptual impact of that path is determined not by its geometry in $\mathcal{S}$ but by its image under a rendering-and-perception map $\Pi : \mathcal{S} \to \mathcal{P}$ into a perceptual manifold $\mathcal{P}$ whose coordinates are attentional focus, revealed narrative information, occlusion structure, emotional intensity, and compositional balance.

Our central result—the geodesic principle of cinematic naturalness—states that cinematographically natural camera movements correspond to paths that locally minimize the perceptual action functional on $\mathcal{P}$, and that the metric governing this functional is the Fisher–Rao metric on the statistical manifold of viewer attentional distributions. Narrative intent is encoded as a potential function that forces paths toward perceptually salient states, yielding a variational equation that unifies classical editing rules (the 180-degree rule, match cuts, continuity editing, montage) as geometric constraints in $\mathcal{P}$.

We then demonstrate that the same variational structure governs painting strokes guided by image-gradient fields, children's drawing as a sequential grammar of spatial action, QWERTY swipe gestures as a trajectory language over a discrete keyboard manifold, and eye saccades as perceptual sampling paths. The topology of gesture classes provides additional invariants that explain why swipe keyboards tolerate large motor variability. We further argue that artistic development constitutes a progressive shift from object-graph representations to field-based and finally radiance-model representations of the same canvas. The framework connects cinematography, game-engine simulation, information geometry, embodied cognition, and trajectory-based generative models.

ÌƖ77L77ƽ

1 Introduction

Many cultural activities that seem superficially unrelated share a common deep structure: they are all instances of directed motion through a structured spatial or perceptual field, where meaning or aesthetic quality emerges from the shape of that motion rather than from its endpoints alone. A filmmaker choosing a camera angle, a painter laying down a brushstroke, a child completing a figure on a page, a typist swiping a word, and a viewer's eye jumping between salient features of a scene are all performing versions of the same cognitive and bodily operation.

This paper proposes a unified mathematical framework for these activities. The central object is a trajectory: a continuous path $\gamma(t)$ through some structured space, where the space is endowed with a geometry that determines which paths are natural, smooth, or costly. We argue that the appropriate notion of naturalness is always determined by a perceptual metric—a metric induced on a manifold of expressive states—rather than by a purely physical or symbolic metric.

The framework has five main components.

(1) Scenario manifold $\mathcal{S}$: the product of camera, actor, lighting, and environment configuration spaces.

(2) Perceptual manifold $\mathcal{P}$: the space of viewer experiential states, carrying a metric derived from information geometry.

(3) Rendering-and-perception map $\Pi : \mathcal{S} \to \mathcal{P}$: the operator that converts scenario states into perceptual states.

(4) Geodesic principle: cinematographic naturalness corresponds to approximate geodesic motion in $\mathcal{P}$ subject to a narrative potential function.

(5) Generality: the same variational structure governs painting strokes, drawing grammars, swipe gestures, eye saccades, and musical perception.

Modern film production already incorporates part of this structure. Virtual production pipelines such as Unreal Engine represent a scene as a game state $g = (C, A, L, E)$ whose components are camera pose, actor configurations, lighting, and environment, and the camera path through that state space is literally a trajectory. But a film rendered with a physical camera also constitutes a path through a latent scenario space, whether or not that space is ever explicitly simulated. The theoretical reinterpretation proposed here treats this implicit structure as the primary object of analysis.

Section 2 defines scenario space formally. Section 3 introduces the perceptual manifold and its metric. Section 4 states the central theorem and derives the variational equations. Section 5 reinterprets classical editing conventions. Section 6 grounds the metric in information geometry. Section 7 extends the framework to painting and visual gradients. Section 8 addresses children's drawing and the grammar of spatial action. Section 9 treats QWERTY swipe gestures as a trajectory language. Section 10 develops topological invariants of gesture classes. Section 11 connects the framework to eye saccades and embodied cognition. Section 12 makes explicit the equivalence between painting traces and other trajectory systems. Section 13 analyzes curvature control in cinematic motion. Section 14 describes the developmental arc from object grammar to radiance model. Section 15 discusses implications for generative models. Section 16 concludes.

2 Scenario Space

2.1 Product Structure

The scenario manifold is the product

$$\mathcal{S} = \mathcal{C} \times \mathcal{A} \times \mathcal{L} \times \mathcal{E},$$

where $\mathcal{C}$ is the camera configuration space, $\mathcal{A}$ is the actor configuration space, $\mathcal{L}$ is the lighting space, and $\mathcal{E}$ is the environmental state space. Each factor is a smooth manifold, so $\mathcal{S}$ inherits a product smooth structure.

Definition 2.1 (Scenario state). A scenario state is a tuple

$$s = (C, A, L, E) \in \mathcal{S}$$

where:

- $C = (x, y, z, \theta, \phi, f) \in \mathcal{C} \cong \mathbb{R}^3 \times S^2 \times \mathbb{R}_{>0}$ encodes camera position, orientation, and focal length;
- $A = (a_1, \ldots, a_k) \in \mathcal{A}$ encodes the pose and animation state of each of k actors, with each a_i in a body-pose Lie group;
- $L \in \mathcal{L}$ encodes lighting configuration (source positions, intensities, color temperatures);
- $E \in \mathcal{E}$ encodes environmental parameters (weather, time of day, set geometry).

For discrete categorical parameters (shot type, costume change) the smooth structure is replaced by a stratified space. We work in the smooth case throughout.

2.2 Films as Paths

Definition 2.2 (Film). A film is a piecewise-smooth path

$$\gamma : [0, T] \to \mathcal{S}.$$

Each frame corresponds to a sample $s_t = \gamma(t)$. The observable image at time t is the output of a rendering operator:

$$I_t = R(s_t), \qquad R : \mathcal{S} \to L^2(\Omega, \mathbb{R}^3),$$

where $\Omega \subset \mathbb{R}^2$ is the image plane and $L^2(\Omega, \mathbb{R}^3)$ is the space of square-integrable color-valued functions on Ω. The observable film is the family $(I_t)_{t\in[0,T]}$, i.e., the image of the trajectory under R.

Remark 2.3. Writing $R : \mathcal{S} \to L^2(\Omega, \mathbb{R}^3)$ makes explicit that an image is an element of an infinite-dimensional function space, not merely a finite array. In a real-time game engine, $\mathcal{S}$ is the engine's world state and R is the rendering pipeline. A film recorded on a physical set has the same abstract structure, with R replaced by the physical optics of the camera and the photochemical or electronic response of the sensor.

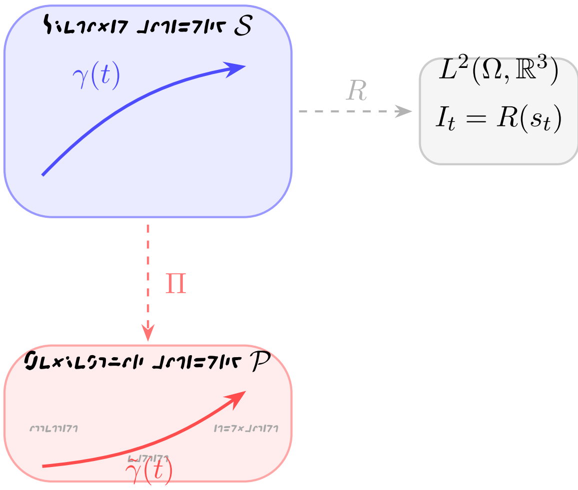


Figure 1: The scenario manifold $\mathcal{S}$, rendering map R, and perceptual manifold $\mathcal{P}$. A film is a path $\gamma(t)$ in $\mathcal{S}$; its aesthetic quality is determined by the induced perceptual path $\tilde{\gamma}(t)$. Selected coordinate axes of $\mathcal{P}$ are indicated inside the lower blob.

3 The Perceptual Manifold

The scenario manifold describes what the world is; the perceptual manifold describes what a viewer experiences. A physically simple camera move can feel disorienting, while a physically elaborate tracking shot can feel effortless. The difference lives not in $\mathcal{S}$ but in $\mathcal{P}$.

3.1 Perceptual Coordinates

Definition 3.1 (Perceptual state). A perceptual state is a point

$$p = (a, r, o, e, q) \in \mathcal{P} \subset \Delta(\mathcal{X}) \times [0,1]^4$$

where $\Delta(\mathcal{X})$ is the probability simplex over the set $\mathcal{X}$ of salient scene objects, and:

- $a \in \Delta(\mathcal{X})$ is the viewer's attentional distribution over salient objects;

- $r \in [0,1]$ is revealed narrative information (cumulative proportion of withheld story information disclosed);
- $o \in [0,1]$ is occlusion coherence (stability of depth ordering and figure-ground relations);
- $e \in [0,1]$ is emotional intensity (arousal level induced by the scene);
- $q \in [0,1]$ is compositional balance (degree of visual equilibrium in the frame).

Treating a as a probability distribution rather than a scalar makes the embedding $\mathcal{P} \subset \Delta(\mathcal{X}) \times [0,1]^4$ natural and connects immediately to the Fisher–Rao metric developed in Section 6.

3.2 Riemannian metric and perception map

The perceptual manifold $\mathcal{P}$ carries a Riemannian metric tensor $g_{ij}(p)$ that measures the cognitive cost of moving between perceptual states. The infinitesimal perceptual distance is

$$ds^2 = \sum_{i,j} g_{ij}(p)\, dp^i\, dp^j.$$

Definition 3.2 (Rendering-and-perception map). The rendering-and-perception map is a smooth map

$$\Pi : \mathcal{S} \to \mathcal{P}$$

that assigns to each screen state the perceptual state it induces in a standard viewer. A film trajectory $\gamma(t)$ in $\mathcal{S}$ induces the perceptual trajectory

$$\tilde{\gamma}(t) = \Pi(\gamma(t)) \in \mathcal{P}.$$

The metric on $\mathcal{P}$ pulls back to a metric on $\mathcal{S}$ via the Jacobian $D\Pi_s : T_s\mathcal{S} \to T_{\Pi(s)}\mathcal{P}$:

$$(D\Pi^* g)_s(u, v) = g_{\Pi(s)}(D\Pi_s \cdot u,\ D\Pi_s \cdot v), \qquad u, v \in T_s\mathcal{S}.$$

In this pullback metric, screen directions that induce large perceptual change are long even if they are physically small, giving a precise sense in which Π compresses or stretches directions according to their perceptual significance.

4 Naturalness as a Geodesic Problem

4.1 Central theorem

Theorem 4.1 (Geodesic Principle of Cinematic Naturalness). Let $\gamma : [0, T] \to \mathcal{S}$ be a screen trajectory and $\tilde{\gamma} = \Pi \circ \gamma$ the induced perceptual trajectory in $(\mathcal{P}, g)$. Among all screen trajectories connecting fixed perceptual endpoints $\tilde{\gamma}(0)$ and $\tilde{\gamma}(T)$, the subjectively natural

[illegible]

$$\mathcal{A}[\tilde{\gamma}] = \frac{1}{2}\int_0^T g_{ij}(\tilde{\gamma}(t))\,\dot{\tilde{\gamma}}^i(t)\,\dot{\tilde{\gamma}}^j(t)\,dt.$$

[illegible]

$$\ddot{\tilde{\gamma}}^k + \Gamma^k_{ij}\,\dot{\tilde{\gamma}}^i\,\dot{\tilde{\gamma}}^j = 0,$$

[illegible] Γ^k_{ij} [illegible] g.

[illegible] $\mathcal{A}$ [illegible] $\mathcal{L}(p,\dot{p}) = \frac{1}{2}g_{ij}(p)\dot{p}^i\dot{p}^j$. [illegible] $\mathcal{A}$ [illegible]

$$\frac{d}{dt}\big(g_{k\ell}\dot{\tilde{\gamma}}^\ell\big) = \frac{1}{2}(\partial_k g_{ij})\dot{\tilde{\gamma}}^i\dot{\tilde{\gamma}}^j,$$

[illegible] $g^{k\ell}$ [illegible] $\Gamma^k_{ij} = \frac{1}{2}g^{k\ell}(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij})$ [illegible]. □

[illegible]

4.2 [illegible]

[illegible] $\mathcal{P}$ [illegible] $V : \mathcal{P} \to \mathbb{R}$, [illegible] V [illegible]

[illegible] 4.2 [illegible]

$$\mathcal{J}[\tilde{\gamma}] = \int_0^T \left[\frac{1}{2}\,g_{ij}(\tilde{\gamma})\,\dot{\tilde{\gamma}}^i\,\dot{\tilde{\gamma}}^j + \lambda\,V(\tilde{\gamma})\right]dt,$$

[illegible] $\lambda > 0$ [illegible]

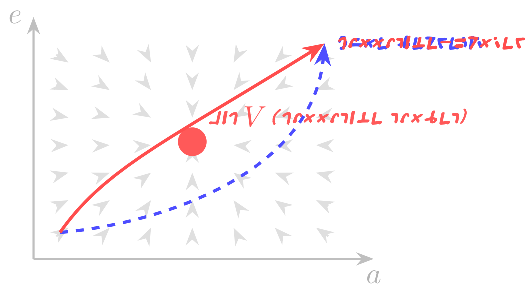
[illegible] 4.3. [illegible] $\mathcal{J}$ [illegible]

$$\ddot{\tilde{\gamma}}^k + \Gamma^k_{ij}\,\dot{\tilde{\gamma}}^i\,\dot{\tilde{\gamma}}^j + \lambda\,g^{k\ell}\,\partial_\ell V(\tilde{\gamma}) = 0,$$

[illegible]

$$\frac{D\dot{\tilde{\gamma}}}{dt} = -\lambda\,\nabla_g V(\tilde{\gamma}).$$

XLJrx+ 4.4. ΊΤL 5rxrJLΊLx λ Lΐ57rLs rlxLΐ17xlrlı ΐ1ıılıL. ʃ rlxLΐ17x ΛΤ7 5xlΛLs ΐ7J57ΐlΊl7lrlı =lı=lrlıı 75LxrΊLs r1 ΐJrlılı λ (Ίrx+7⊥ΐ+ıı, 757⁼lΐ); r rlxLΐ17x ΛΤ7 5xl7xlΊΛLs Ίrxxr1⊥L l1=7xJr1l71 rLΊΐlΊıı 75LxrΊLs r1 lırx+L λ (LlΐLΊΐΊLl1, +xLLΊ+xrΐΐ).



=l+=xL 2: Ίrxxr1⊥L =7xΐl1+ ΐL1rΐ ΊΤL 5LxΐL5Ί=rlı 5r1Τ Ί7Λrxr ΊΤL r11xrΐ17x Jl1 V (ΊΤL Ίrxxr1⊥Llıı lJ57xΊrΊ1 ΐΊrΊL) r1 ΊΤL ΐ7ΐΊ 7= ΐ7JL ΐJ77ΊΤLΐΐ. Ί77xrlΊrΊL r/Lΐ a (r11LΊΊl7Ίrlı =7ΐ=ΐ) r1r e (LJ7Ίl7Ίrlı l11LΊΐlΊıı) lırΐLlı 1Λ7 rlJLΊΐl71ΐ 7= $\mathcal{P}$.

5 Lrlıl1+ rΐ Ίxr≷LΐΊ7xıı Ίxr1ΐ=7xJr1l71

=llıJ Lrlıl1+ l11x7r=ΐLΐ rlΐΐ71Ίl1=lılLΐ 7x ΐ1x=ΐΊ=xLr Ίxr1ΐ=7xJr1l71ΐ l1 ΊΤL ΐΐL1rxl7 Ίxr≷LΐΊ7xıı. ΛL l11Lx5xLΊ ΐlırΐΐlrlı Lrlıl1+ 75Lxr1l71ΐ rΐ +L7JLΊxlr 75Lxr1l71ΐ 71 5rΊΤΐ l1 $\mathcal{P}$, rxrΛl1+ 71 ΊΤL ΊΤL7xLΊlΐrlı =xrJLΛ7x+ 7= Ј7xrΛLlılı r1r Ί7Ί7J5ΐ71 [4] =7x ΊΤL =ΊrLxlıılı1+ =llıJ–lır1+=r+L Ί1/7Ί7Jıı.

5.1 Ί=Ίΐ r1r 5LxΐL5Ί=rlı rlΐΊrΊΐL

ʃ Τrxr Ί=Ί r1 ΊlJL t_0 lΐ r ≷=J5 rlΐΐ71Ίl1=lΊıı: $\gamma(t_0^-) \neq \gamma(t_0^+)$. l1 $\mathcal{P}$ ΊΤL ΐ7ΐΊ 7= ΊΤL Ί=Ί lΐ 5x757xΊl71rlı Ί7 $d_{\mathcal{P}}(\tilde{\gamma}(t_0^-), \tilde{\gamma}(t_0^+))$. ʃ Ί=Ί lΐ l1⊥lΐlΐlıL ΛΤL1 ΊΤlΐ rlΐΊrΊΐL lΐ ΐJrlılı, l.L., ΛΤL1 ΊΤL Ί=Ί 5xLΐLx⊥Lΐ r11LΊΊl7Ίrlı =7ΐ=ΐ a, 7ΐΐlı=ΐl71 ΐΊΤLxLΊΐL o, r1r ΐ7J57ΐlΊl7Ίrlı ΐlırΊΐL q. ΊΤlΐ lΐ ΊΤL +L7JLΊxlΐ ΐ7Ί1L1 7= ΐ7Ί1l1=lΊıı Lrlıl1+.

5.2 Γlΐΐ7lı⊥Lΐ r1r Λl5Lΐ

ʃ rlΐΐ7lı⊥L ΐlıL1rΐ Ί1Λ7 ΐΐL1rxl7 ΐΊrΊLΐ:

$$I_t = (1 - \alpha(t))\, R(\gamma_1(t)) + \alpha(t)\, R(\gamma_2(t)), \quad \alpha : [0,1] \to [0,1].$$

ΊΤL rlΐΐ7lı⊥L lΐ 5LxΐL5Ί=rlılıı ΐJ77ΊΤ ΛΤL1 ΊΤL ΐ7=xΊL 5rΊΤΐ $\Pi(\gamma_1)$ r1r $\Pi(\gamma_2)$ rxL ΐlı7ΐL l1 $\mathcal{P}$.

5.3 J771r+L

ʃ J771r+L ΐLJ=L1ΊL lΐ r 5lLΊLΛlΐL Ίxr≷LΐΊ7xıı $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_k)$ Λl1Τ ≷=J5 rlΐΐ71Ίl1=lΊlLΐ r1 ΊΤL ≷71Ίΐ. J771r+L rLlılΐLxrΊLlıı ΐxLr1Lΐ lırx+L 5LxΐL5Ί=rlı ≷=J5ΐ, ΐ=Ί lΊLrlılı rlΊ7+ ΐLJr11lΊrlılıı meaningfully axes (primarily e and r), so that perceptual cost is compensated by conceptual gain. Eisenstein's principle of intellectual montage is precisely that large excursions in $\mathcal{P}$ along the r-axis (narrative revelation) can be aesthetically productive even at high perceptual cost in the a and o coordinates.

5.4 The 180-degree rule as a chart condition

Proposition 5.1. The 180-degree rule is the condition that the camera trajectory remains within a single local chart of $\mathcal{P}$ with respect to the orientation coordinate o. Crossing the line forces a coordinate singularity in that chart.

Proof. If actors lie along an axis $\ell \subset \mathbb{R}^3$, a camera on one side of ℓ assigns consistent left-right screen positions to the actors. Crossing ℓ reverses this assignment, corresponding to an abrupt change in o that cannot be smoothed by a path of small g-length in $\mathcal{P}$. Formally, the orientation coordinate o has a fold singularity at the axial plane, so the camera trajectory must remain in one connected component of $\{o > 0\}$ to remain within a single chart. □

5.5 Match cuts and eyeline matches

A match cut at time t_0 satisfies $\|\tilde{\gamma}_1(t_0^-) - \tilde{\gamma}_2(t_0^+)\|_{\mathcal{P}} < \varepsilon$ for small ε. The match cut is the editorial operation that minimizes perceptual distance across a cut. An eyeline match aligns the attention distribution a of one shot with the implied object of attention in the next, preserving continuity in the $\Delta(\mathcal{X})$ component of $\mathcal{P}$.

6 Information Geometry of the Perceptual Metric

6.1 Viewer attention as a statistical manifold

We ground the metric g on $\mathcal{P}$ in information geometry by modeling the viewer's attentional coordinate a as a probability distribution $P_t \in \Delta(\mathcal{X})$ over salient scene objects. When P_t is parameterized by $\theta \in \Theta$, the Fisher information metric is

$$g_{ij}^F(\theta) = \mathbb{E}_{P(\cdot;\theta)}\left[\frac{\partial \log P(x;\theta)}{\partial \theta^i}\frac{\partial \log P(x;\theta)}{\partial \theta^j}\right].$$

This is the unique (up to scale) Riemannian metric on $\Delta(\mathcal{X})$ invariant under sufficient statistics [5, 1].

Theorem 6.1. The perceptual metric g on $\mathcal{P}$ can be taken to dominate the Fisher-Rao metric on the attentional component $\Delta(\mathcal{X})$. A scenario trajectory γ is perceptually smooth if and only if the induced attentional path $t \mapsto P_t$ has bounded Fisher-Rao velocity.

[illegible] P_t [illegible] a[illegible] $\int_0^T g_{ij}^F(P_t)\dot{P}_t^i\dot{P}_t^j\,dt < \infty$, [illegible] $\Delta(\mathcal{X}) \hookrightarrow \mathcal{P}$, [illegible] $\mathcal{A}[\tilde{\gamma}]$. $\square$

6.2 [illegible]

[illegible] P^- [illegible] P^+ [illegible]

$$W_2(P^-, P^+) = \inf_{\pi\in\Pi(P^-,P^+)} \left(\int_{\mathcal{X}^2} d(x,y)^2 \, d\pi(x,y) \right)^{1/2},$$

[illegible] π [illegible] P^- [illegible] P^+ [15]. [illegible] $W_2(P^-, P^+)$ [illegible]

[illegible] 6.2. [illegible]

[illegible]

7 [illegible]

7.1 [illegible]

[illegible] $B_t : \Omega \to \mathbb{R}$, [illegible] $\Omega \subset \mathbb{R}^2$ [illegible] $\gamma_{[illegible]} : [0, \tau] \to \Omega$ [illegible]

$$B_{t+dt}(x,y) = B_t(x,y) + \int_0^{dt} K(x - \gamma_x(t),\, y - \gamma_y(t))\, dt,$$

[illegible] K [illegible]

7.2 [illegible]

[illegible] $\nabla I = (I_x, I_y)$ [illegible] $I : \Omega \to \mathbb{R}$, [illegible] $G_x = S_x * I$, $G_y = S_y * I$, $|\nabla I| = \sqrt{G_x^2 + G_y^2}$. [illegible] $T = (-I_y, I_x)$, tangent to level curves of I, yielding the stroke differential equation

$$\frac{d\gamma_{\text{paint}}}{dt} = \alpha\, T(\gamma_{\text{paint}}(t)) + \beta\, \nabla I(\gamma_{\text{paint}}(t)).$$

This has the same form as the narratively forced geodesic of Section 4: motion follows a vector field derived from perceptual structure, balancing contour-following (α) against gradient-crossing (β) in a manner precisely analogous to the balance between perceptual smoothness and narrative urgency in cinematography.

7.3 Structure Tensor and Dominant Orientation

The structure tensor

$$J = \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$$

summarizes local orientation. Its dominant eigenvector gives the preferred direction of edges and forms [13, 9].

Proposition 7.1. A painter who aligns brushstrokes with the dominant eigenvector field of J is integrating a trajectory through the metric induced by J on Ω. Each stroke is a geodesic segment in this derived metric.

Painterly rendering algorithms such as those of Hertzmann [9] make this explicit computationally, generating curved brushstrokes aligned with image curvature flow—a direct algorithmic implementation of the gradient-field integration described here.

8 Drawing as a Sequential Grammar of Spatial Action

8.1 Goodnow's Grammar of Action

Jacqueline Goodnow's empirical studies demonstrate that children's drawings are not arbitrary marks but records of structured action sequences [8]. Children follow systematic strategies for initiating figures, attaching new elements, and expanding outward—a procedural grammar of action governing how gestures combine to produce spatial structures.

In the present framework, a drawing is a sequence of trajectories $\gamma_1, \gamma_2, \ldots, \gamma_n$ each depositing a stroke on the canvas field B, governed by an iterative mapping

$$B_{t+1} = F(B_t, a_t),$$

where a_t is the stroke action at step t. Goodnow's attachment and expansion rules correspond to F: new strokes are placed near existing structures, extend boundaries, or occupy regions that preserve compositional balance.

8.2 [illegible]

[illegible]

8.3 [illegible]

[illegible] [10] [illegible]

[illegible] [12] [illegible]

9 [illegible]

9.1 [illegible]

[illegible] $K = (V, E)$ [illegible] $\phi : V \to \mathbb{R}^2$. [illegible] K [illegible]

[illegible] 9.1 [illegible] $w = w_1 \cdots w_n \in \Sigma^*$, [illegible] $P(w) = (\phi(w_1), \phi(w_2), \ldots, \phi(w_n)) \subset \mathbb{R}^2$.

[illegible] 9.2 [illegible] $P : \Sigma^* \to \mathcal{T}$ [illegible] $(\Sigma^*, \cdot)$ [illegible] $(\mathcal{T}, \cdot)$ [illegible]

[illegible] $w = uv$ [illegible] $P(w) = P(u) \cdot P(v)$ [illegible] $P(u)$'s [illegible] □

9.2 Geometric invariants

Each word w has a geometric signature:

$$
\begin{aligned}
L(w) &= \textstyle\sum_{i=1}^{n-1} \|\phi(w_{i+1}) - \phi(w_i)\| && \text{(path length)},\\
C(w) &= \textstyle\sum_{i=2}^{n-1} |\theta_i| && \text{(turn complexity)},\\
X(w) &= \text{(self-intersection count)},\\
\kappa_{\text{close}}(w) &= 1 - \frac{\|\phi(w_n)-\phi(w_1)\|}{L(w)} && \text{(closure index)}.
\end{aligned}
$$

Definition 9.3 (Trajectory equivalence). Words $w, u \in \Sigma^*$ are trajectory-equivalent if $d_F(\widetilde{P}(w), \widetilde{P}(u)) < \varepsilon$, where $\widetilde{P}$ is the arc-length-normalized version of P and d_F is Fréchet distance.

9.3 Geometric taxonomy of the lexicon

Shape class	Example	Dominant invariant
Loop / circle	tent	high κ_{close}
Rectilinear	route	few turn directions
Braided	magic	$X(w) \geq 1$
Arch	con	smooth convex arc
Hook	do	large single turn, short L
Ribbon sweep	all	large L, small C
Spiral	crying	increasing radius
Zigzag	zeno	alternating diagonal segments

Corollary 9.4. Any lexicon $\mathcal{L} \subset \Sigma^*$ inherits a trajectory semantics via P, making it simultaneously a symbolic language and a geometric codebook. The natural-language vocabulary occupies a proper submanifold of the full swipe trajectory space.

The swipe keyboard thus corresponds to the painter's canvas in discretized form: instead of depositing pigment continuously, the gesture snaps to discrete letter centers. The dual reading of a swipe word is $w \mapsto \big(\text{semantic meaning, gesture class}\big)$. This was first formally computationally in the SHARK^2 system of Kristensson and Zhai [11], which demonstrated that large-vocabulary shorthand writing can be recovered from gesture trajectories alone.

10 Topology of Gesture Classes

Trajectory languages can be analyzed through topological invariants preserved under continuous deformation without crossing a forbidden region.

Definition 10.1 (Topological equivalence). Two trajectories $\gamma_1, \gamma_2 \in \mathcal{T}$ are topologically equivalent if one can be continuously deformed into the other within $\mathbb{R}^2 \setminus \mathcal{O}$, where $\mathcal{O}$ is the set of obstacle regions (keyboard edges, object boundaries).

Key invariants include:

$$X(\gamma) = \text{signed self-intersection count},$$
$$\omega(\gamma) = \text{winding number around a reference point},$$
$$\kappa_{\text{sign}}(\gamma) = \text{signed curvature sequence}.$$

These partition trajectory space into gesture classes:

Class	Topological invariant	Example
Loop	$\omega = 1$	circular swipe or closed brushstroke
Spiral	$\omega > 1$	expanding trajectory
Zigzag	alternating curvature sign	segmented gesture
Pretzel	$X \geq 1$	self-crossing trajectory

The significance is that human motor systems naturally produce gestures that remain stable within a topological class: small variations in speed or curvature do not change the qualitative shape. Swipe keyboards tolerate large motor variability precisely because recognition operates on topological class membership rather than exact geometric matching. This also explains why artists develop characteristic stroke families that constitute a personal topological vocabulary of movement.

11 Eye Saccades and Embodied Perception

11.1 Saccades as Perceptual Sampling Trajectories

Visual perception does not occur through smooth continuous scanning. The eye moves through sequences of rapid fixation jumps called saccades. If fixation points are $p_1, p_2, \ldots, p_n \in \mathcal{I}$, the visual scan path is a trajectory $\gamma_{\text{eye}}(t)$ through the image plane.

Yarbus [16] showed that scan paths are not random: they are governed by a salience field $\sigma : \mathcal{I} \to \mathbb{R}_{\geq 0}$, where high salience corresponds to edges, faces, motion, and high-contrast regions. The saccade trajectory approximately maximizes accumulated salience subject to a motor cost for large eye movements—a constrained optimization structurally identical to the variational principle of Theorem 4.1.

11.2 Three trajectory systems

Three trajectory systems operate jointly in visual culture:

[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible] $\gamma(t)$ [illegible]		

[illegible]

11.3 [illegible]

[illegible] [6]. [illegible] $m_{t+1} = F(m_t)$, [illegible] [3] [illegible]:

$$a_1 \to a_2 \to \cdots \to a_n \to a_1,$$

[illegible]

[illegible]:

> [illegible]

12 [illegible]

12.1 [illegible]

[illegible] $\gamma_{[illegible]} : [0,T] \to \Omega$ [illegible] $(\gamma) = \{\gamma(t) : t \in [0,T]\}$. [illegible] γ [illegible] $\mathcal{S}$, [illegible]

[illegible]:

[illegible]	[illegible]	[illegible]
[illegible]	[illegible] K	[illegible]
[illegible]	[illegible] B_t ([illegible])	[illegible]
[illegible]	[illegible] σ	[illegible]
[illegible]	[illegible] $\mathcal{S}$	[illegible]
[illegible]: $\gamma(t)$ [illegible] F.		

12.2 [illegible]

[illegible]

12.3 [illegible]

[illegible]

13 [illegible]

[illegible]

[illegible] 13.1 ([illegible]). [illegible] $\tilde{\gamma}(t) = \Pi(\gamma(t))$ [illegible]. [illegible] t [illegible]

$$\kappa(t) = \frac{\|\dot{\tilde{\gamma}}(t) \times \ddot{\tilde{\gamma}}(t)\|}{\|\dot{\tilde{\gamma}}(t)\|^3}.$$

[illegible]

[illegible]	[illegible]
[illegible]	$\kappa(t)$ [illegible]
[illegible]	[illegible] $\kappa(t)$
[illegible]	[illegible]
[illegible]	[illegible] κ [illegible]
[illegible]	κ [illegible] V–[illegible]

[illegible] 13.2. [illegible] $\kappa(t)$. [illegible] κ [illegible] κ [illegible] $\mathcal{P}$ [illegible] Π.

[illegible] $\int_0^T \kappa(t)\,dt$ [illegible] $\mathcal{P}$ [illegible] $\mathcal{S}$ [illegible]

14 [illegible]

[illegible]

[illegible] 1: [illegible] $G = (R, E)$ [illegible] R_i [illegible] E [illegible]

[illegible] 2: [illegible] $I(x, y)$. [illegible]

[illegible] 3: [illegible]

$$L_o(x,\omega) = L_e(x,\omega) + \int_\Omega f_r(x,\omega',\omega)\, L_i(x,\omega')\,(\omega'\cdot n)\, d\omega'.$$

[illegible]

[illegible]

15 [illegible]

[illegible]

[illegible] $\tilde{\gamma}_1, \ldots, \tilde{\gamma}_N$ [illegible]

(a) [illegible] g_θ [illegible] $\mathcal{P}$ [illegible]

(b) [illegible] V_ϕ [illegible]

(c) [illegible] Π_ψ [illegible]

[illegible] $\gamma(t)$ [illegible] $\tilde{\gamma}(t)$ [illegible] $\mathcal{J}[\tilde{\gamma}]$ [illegible]

[illegible] 15.1. [illegible]

[illegible]

16 [illegible]

[illegible]

[illegible] 4.1 [illegible] $\mathcal{P}$, [illegible] V. [illegible] 6.1 [illegible] $\mathcal{P}$; [illegible] 5.1 [illegible]

[illegible]

[1] [illegible], 2000.

[2] [illegible], 1954.

[3] [illegible], 2007.

[4] [illegible], 2013.

[5] [illegible], 1982.

[6] [illegible], 2016.

[7] [illegible], 1992.

[8] [illegible], 1977.

[9] [illegible] 1998. [illegible], 1998, [illegible] 453-460.

[10] [illegible], 1969.

[11] [illegible] 2004. [illegible], 2004, [illegible] 43-52.

[12] [illegible], 1947.

[13] [illegible] 12.7 (1990): 629-639.

[14] [illegible], 2019.

[15] [illegible], 2009.

[16] [illegible], 1967.