

Moiré Video Authentication: A Physical Signature Against AI Video Generation

Yuan Qing*, Kunyu Zheng*, Lingxiao Li*, Boqing Gong, and Chang Xiao**

Boston University, Boston, USA

{ymqing, kunyuz, lxli, bgong, xchang}@bu.edu

https://yuanqing-ai.github.io/physical_video_signature/

Abstract. Recent advances in video generation have made AI-synthesized content increasingly difficult to distinguish from real footage. We propose a physics-based authentication signature that real cameras produce naturally, but that generative models cannot faithfully reproduce. Our approach exploits the Moiré effect: the interference fringes formed when a camera views a compact two-layer grating structure. We utilize the *Moiré motion invariant* as a verification criterion: under real image formation, fringe phase and grating image displacement are linearly coupled by optical geometry, independent of viewing distance and grating structure. A verifier extracts both signals from video and tests their correlation. We validate the invariant on both real-captured and AI-generated videos from multiple state-of-the-art generators, and find that real and AI-generated videos produce significantly different correlation signatures, suggesting a robust means of differentiating them. Our work demonstrates that deterministic optical phenomena can serve as physically grounded, verifiable signatures against AI-generated video.

Keywords: Video Authentication · Moiré Pattern · Video Watermarking · Physical-Based Forensics

1 Introduction

In February 2024, a finance worker at a multinational firm was tricked into transferring \$25 million after attending a video call in which every other participant, including the company’s chief financial officer, were impersonated by AI-generated video [8]. In October 2025, an AI-generated video mimicking a news broadcast falsely depicted an Irish presidential candidate announcing her withdrawal from the race, spreading across social media for hours before being removed [2]. Spreading unauthenticated misinformation by generative AI is no longer a hypothetical scenario; it is the new reality. Generative video models such as Sora [5], Runway [34], and Seedance [6] now produce footage so convincing that human viewers, automated detectors, and even forensic analysts struggle to distinguish it from real recordings [10]. As AI-generated video increasingly

* Equal contribution

** Corresponding author

spreads misinformation and erodes public trust, the need for a reliable, theoretically grounded mechanism to prove that a video is *authentic* has become more urgent than ever.

Existing AI video detection broadly fall into two categories. *Post-hoc* forensic detectors analyze pixel-level artifacts, temporal inconsistencies, or learned statistical fingerprints to flag AI-generated content [17, 33, 41]. While effective against earlier generators, these methods are locked in an arms race: each new model eliminates the artifacts that detectors exploit, demanding constant retraining and offering limited lasting guarantees. *Digital watermarking* approaches, such as C2PA metadata [9] and SynthID [15], embed provenance information into media at creation time. However, metadata can be stripped, re-encoded, or forged, and digital watermarks degrade under common post-processing operations. Crucially, neither paradigm provides an unforgeable physical link between the recorded scene and the resulting video.

We propose a fundamentally different approach: a **physical authentication signature**. A real video is produced by an optical system governed by physical laws, whereas generative models learn statistical correlations from data and rarely model the underlying physics. A physical phenomenon whose appearance is tightly coupled to precise optical geometry therefore serves as a signature that real cameras produce naturally but that current video generative models cannot precisely reproduce.

We instantiate this idea through the Moiré effect [1], an intriguing and widely observed yet often overlooked optical phenomenon. Moiré effect are the interference fringes that appear when two fine periodic structures, such as line gratings, are overlaid (Fig. 1-a). These fringes are highly sensitive to the precise geometric relationship between the camera and the grating assembly: as the camera moves, the fringes shift according to deterministic optical laws (Fig. 1-b). This tight coupling between camera motion and fringe motion forms a physics-based signature that real optical systems produce naturally (Fig. 1-c) but that current video generation models cannot reliably synthesize (Fig. 1-d).

The core assumption underlying our approach is that current video generation models are not grounded in physical simulation. They learn to produce plausible-looking frames from large-scale data, but they do not model the wave-optics interactions that govern Moiré fringe formation and evolution. Accurately reproducing the effect would require solving the exact optical problem for every frame: tiny shifts in camera position produce measurable fringe displacement, and the mapping from camera motion to fringe motion follows a strict mathematical relationship derived from first principles [1]. This stands in sharp contrast to the statistical interpolation that current generative models perform. Even if future models were trained on abundant footage containing Moiré patterns, learning the visual appearance of fringes can hardly, if not impossibly, equip a model to solve the underlying optical equations.

Leveraging this insight, we propose a practical authentication system that works as follows. A compact grating assembly, consisting of a lenticular sheet mounted above a printed stripe pattern, is placed in the scene (*e.g.*, worn by a



Fig. 1: (a) The Moiré effect, created by overlaying two repetitive layers. (b) Our prototype Moiré signature assembly. Viewing the assembly from slightly different camera angles (left vs. right) causes the fringes to shift in phase according to deterministic physical laws. (c) In a real video, Moiré fringes appear naturally and shift predictably with camera movement. (d) In an AI-generated video, the fringes distort, and their phase shifts do not adhere to the underlying physics.

speaker at a press conference, see Fig. 4). Given a recorded video with relative movement between the assembly and the camera (which occurs naturally from the speaker’s body motion even when the camera is static) our algorithm (1) extracts the Moiré fringe pattern and maps it to a canonical coordinate frame, (2) tracks the fringe phase changes over time, and (3) computes the correlation between the observed fringe phase and the displacement of the assembly in the video. A high correlation certifies authenticity; a low or absent correlation flags the video as potentially synthetic. Importantly, analogous to digital watermarking, our method provides positive authentication for videos that contain the signature; it does not make claims about videos lacking one.

This paper contributes a novel idea supported by theoretical grounding and empirical evidence for validating the feasibility. Our contributions are:

- **A novel Moiré-based video authentication framework.** We introduce the concept of using the Moiré effect as a passive, physics-based authentication signature. We utilize the Moiré motion invariant as a verification criterion: fringe phase displacement and grating image displacement are linearly coupled; the resulting correlation is independent of viewing position and grating structure, yielding a verifiable signal extractable from video alone.
- **A practical implementation and feasibility validation.** We build a proof-of-concept system using off-the-shelf materials (a lenticular sheet, printed patterns, and ArUco markers) and demonstrate that the physics-predicted correlation holds in real captured video, confirming the practical viability of the approach.
- **A threat model analysis.** We systematically examine how our method performs against a spectrum of attack scenarios, characterizing the conditions under which the Moiré signature remains robust. We also synthesize

a large number of AI-generated videos in our best effort to reproduce the Moiré pattern, and the results show that even cherry-picked and carefully engineered generated videos fail to achieve correlation comparable to that of real video.

2 The Moiré Effect

A Moiré pattern is a large-scale interference pattern that emerges when two fine periodic structures are superimposed with a slight difference in their spatial frequencies [1]. Moiré patterns appear ubiquitously in everyday life, from the shimmering bands seen through overlapping fences to the rippling fringes visible when a digital display is photographed at certain positions, and they have found broad application in optical metrology [7, 14, 43], mechanical deformation analysis [1, 23, 36, 39], and computational art [20, 37].

In our system, the Moiré pattern is generated by a compact structure consisting of two layers of line gratings with slightly different spatial periods, separated by a thin gap of thickness g . Because the gap causes each viewing position to reveal a different relative alignment of the two layers, the superposition produces prominent Moiré fringes whose positions are extremely sensitive to camera-grating geometry. The fringes act as a *geometric amplifier*: a microscopic change in viewing position produces a macroscopic fringe shift, yielding a visually prominent response that is easily captured on video yet difficult for generative models to synthesize correctly.

Without loss of generality, we restrict our analysis to **one-dimensional Moiré patterns** (parallel line fringes), which is the configuration used in our implementation. However, two-dimensional Moiré patterns formed by crossed or rotated gratings [1, 43] can be easily implemented and extended in the same principle. As a feasibility study, in this paper we only consider the 1D scenario.

2.1 Theory

Our optical derivation builds on classical Moiré theory, as established in prior textbooks [1, 23, 39]. We illustrate the relationship between camera motion and Moiré fringe displacement for the one-dimensional two-layer grating structure. Consider two aligned gratings with spatial periods p_f (front) and p_r (rear). From standard Moiré theory, their superposition produces fringes with beat period $p_m = p_f p_r / |p_f - p_r|$. A lateral relative shift δ between the layers displaces the fringes by $|\Delta x_m| = M|\delta|$, where $M = p_f / |p_f - p_r|$ is the **Moiré magnification factor**.

In our structure the layers are rigidly fixed; the apparent shift δ arises from viewing-angle parallax. We analyze motion along the axis x perpendicular to the grating lines (the only direction that produces fringe shifts). A camera at distance D undergoing a lateral displacement Δx_c sees an apparent layer shift $\delta = g \Delta x_c / D$, where g denotes the gap between the two grating layers. Converting fringe displacement to **unwrapped phase** ($\Delta\phi = 2\pi\Delta x_m / p_m = 2\pi\delta / p_r$,

where the p_f and $|p_f - p_r|$ terms cancel) and substituting δ yields

$$\Delta\phi = \pm \frac{2\pi g}{p_r \cdot D} \cdot \left(1 + \frac{g}{D}\right) \cdot \Delta x_c, \quad (1)$$

where the sign depends on coordinate convention and the factor $(1 + g/D)$ accounts for the rear grating's projected period on the front grating plane.

2.2 Our Insight: The Correlation Invariant

The preceding relationship shows that Moiré fringe displacement depends on camera motion. We now show that this same quantity governs the apparent motion of the grating structure itself in the camera image, yielding a correlation invariant that is independent of the unknown distance D . We first derive the invariant for pure camera translation, then generalize to arbitrary camera motion that includes rotation.

In our physical setting, the grating gap g is on the order of millimeters while the observation distance $D \geq 0.5$ m, so the factor $(1 + g/D)$ in Eq. (1) is negligible. We therefore approximate Eq. (1) as

$$\Delta\phi = \pm \frac{2\pi g}{p_r \cdot D} \cdot \Delta x_c. \quad (2)$$

Pure translation. Under a pinhole camera model with focal length f , a pure translational camera displacement Δx_c along the perpendicular axis at distance D from the grating structure causes the structure's centroid to shift in the image plane by

$$\Delta u_{\text{trans}} = \pm f \cdot \frac{\Delta x_c}{D},$$

where the sign depends on coordinate convention. Combining this with Eq. (2), we eliminate the common factor $\Delta x_c/D$ to obtain

$$\Delta\phi = \pm \frac{2\pi g}{p_r \cdot f} \cdot \Delta u_{\text{trans}}. \quad (3)$$

General camera motion. In practice, camera motion is a combination of translation and rotation. Under the pinhole model, a small camera rotation $\Delta\theta$ about the axis parallel to the grating lines (*i.e.* pan or tilt that displaces the structure image along the perpendicular axis x) produces an additional image displacement Δu_{rot} . The total observed image displacement of the structure is therefore

$$\Delta u = \Delta u_{\text{trans}} + \Delta u_{\text{rot}}.$$

Critically, rotation does *not* induce parallax between the two grating layers. A pure rotation shifts both layers identically in the image, producing zero apparent relative displacement ($\delta = 0$) and hence zero fringe phase change. Only the translational component Δu_{trans} generates parallax and drives $\Delta\phi$. Therefore Eq. (3) holds in the general case, with $\Delta u_{\text{trans}} = \Delta u - \Delta u_{\text{rot}}$.

For an authentic video, the temporal sequences of fringe phase changes $\{\Delta\phi_t\}$ and translational image displacements $\{\Delta u_{\text{trans},t}\}$ are linearly coupled with near-unity absolute correlation:

$$|\rho| = |\text{Corr}(\{\Delta\phi_t\}, \{\Delta u_{\text{trans},t}\})| \approx 1. \quad (4)$$

We call this the **Moiré motion invariant**. The proportionality constant need not be known to the verifier: D has cancelled entirely, and although f , g , and p_r may be unknown, they are fixed for a given camera and grating structure. Because correlation measures linear coupling regardless of the slope’s value, a verifier simply extracts both sequences from the pixel data and checks whether they are linearly coupled, without requiring any calibration or knowledge of the scene geometry.

In general, isolating Δu_{trans} from the observed image displacement Δu requires estimating and subtracting the rotational component Δu_{rot} , which can be done purely from the video, for instance by tracking distant background features or decomposing camera pose via multi-point correspondence methods such as Perspective-n-Point (PnP) pose estimation [47]. In our implementation, we perform this rotation compensation using the four ArUco fiducial markers already present on the grating assembly as 3D-to-2D correspondences for PnP-based camera pose recovery, allowing us to isolate the translational displacement component. We describe our concrete implementation in Sec. 3.

In an AI-generated video, no physical grating structure exists; the generator must synthesize fringe patterns frame by frame from learned statistics. Because the generator does not solve the optical equation, the synthesized fringe motion will not maintain the strict linear coupling to the structure’s position required by the invariant. Even small deviations, a fringe that shifts too fast, too slow, or in the wrong direction by a fraction of a cycle, will depress $|\rho|$ far below 1. This theoretical gap between physics-governed and statistically generated fringe behavior is the foundation of our authentication algorithm.

3 Proof-of-Concept Implementation

We now describe our proof-of-concept system for validating the Moiré motion invariant. We first detail the physical grating assembly, then present the algorithmic pipeline that extracts both the fringe phase signal and the grating assembly displacement signal from a recorded video and computes their correlation.

3.1 Physical Setup

Our grating assembly consists of three layers bonded together into a flat, rigid structure (see Fig. 2-a):

- **Front layer (viewing grating).** A 1 mm-thick lenticular lens sheet (50 LPI, period $p_f = 0.508$ mm) whose cylindrical lenses act as angular apertures, so that small lateral camera movements change the visible portion of the rear layer and shift the Moiré fringes.

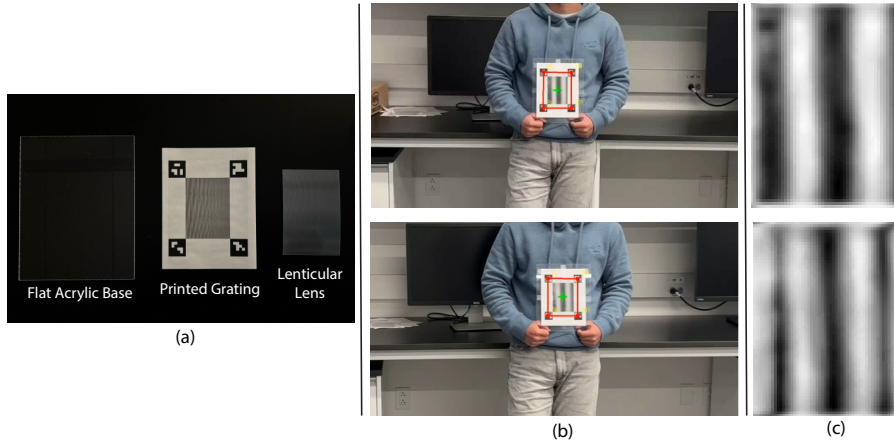


Fig. 2: (a) Exploded view of the grating assembly, showing its three constituent layers: a flat acrylic base, a printed grating, and a lenticular lens. (b) Two distinct video frames illustrating our algorithm detecting the ArUco markers to isolate the Moiré fringe region. (c) The extracted Moiré fringes after a canonical transformation, corresponding to the frames in (b). The phase shift of the fringes between the two frames is clearly visible.

- **Rear layer (reference grating).** A line pattern printed on paper with pitch $p_r = 0.52$ mm. The slight mismatch between p_f and p_r produces Moiré fringes with a beat period of $p_m = p_f \cdot p_r / |p_f - p_r| \approx 22$ mm, yielding a Moiré magnification factor of $M = p_f / |p_f - p_r| \approx 42\times$.
- **Base layer.** A flat acrylic sheet that serves as the structural substrate, holding all layers in alignment.

The three layers are assembled and attached to a flat backing surface. For ease of prototyping, we place four ArUco fiducial markers (from the DICT_4X4_50 dictionary) at the corners of the grating assembly to simplify bounding-box extraction and perspective correction. Any robust localization method could replace them in a production system.

3.2 Verification Pipeline

Given an input video containing the grating assembly, our pipeline extracts two independent temporal signals – the fringe phase and the grating assembly displacement – and tests their linear coupling. The pipeline proceeds in three stages.

Stage 1: Grating Assembly Tracking and Fringe Extraction. We localize the grating assembly in each frame by detecting and tracking the four ArUco corner markers, yielding per-frame corner coordinates and a centroid trajectory $\bar{z}^{(t)}$. Using the four corners, we compute a per-frame homography that warps

the grating assembly region into a canonical, front-parallel rectangle and apply CLAHE [50] to enhance fringe visibility. We then determine the fringe orientation via contour analysis [38], rotate the image to align fringes vertically, and collapse the 2D image into a 1D intensity profile by averaging along the fringe direction. Fig. 2-(b, c) shows the tracked Moiré region and the corresponding warped canonical frames from our pipeline.

Stage 2: Phase Tracking. We extract the fringe phase from the 1D profile using a one-dimensional FFT. Because the profile is approximately sinusoidal at the Moiré beat frequency, the phase of the corresponding Fourier component directly encodes the lateral position of the fringe pattern. On the first frame, we identify the dominant frequency peak and lock its bin k^* ; on each subsequent frame we read the phase at k^* :

$$\phi^{(t)} = \arg \left[\mathcal{F}\{I^{(t)}\}(k^*) \right].$$

To avoid 2π discontinuities, we track the phase incrementally: we wrap each frame-to-frame difference to $[-\pi, \pi]$ using $\text{wrap}(\theta) = ((\theta + \pi) \bmod 2\pi) - \pi$ and accumulate the deltas into a smooth cumulative phase signal:

$$\Phi^{(t)} = \sum_{\tau=1}^t \text{wrap}(\phi^{(\tau)} - \phi^{(\tau-1)}).$$

Stage 3: Displacement Estimation and Correlation. As derived in Sec. 2.2, camera rotation does not induce parallax between the grating layers and therefore does not drive fringe phase change; only the translational component of camera motion does. The raw centroid trajectory $\bar{c}^{(t)}$ from Stage 1 conflates both components, so we decompose it using PnP-based pose estimation. The four ArUco marker centers provide four coplanar 3D-to-2D point correspondences per frame, which we pass to a PnP solver (*i.e.*, OpenCV’s `cv::solvePnP`) to recover the camera pose (R_i, t_i) for each frame. We use a pinhole camera model with approximate intrinsics computed from image size ($f_x = f_y = \alpha \max(W, H)$). We then isolate the translation-only displacement by re-projecting the 3D centroid of the marker layout using each frame’s camera position but the reference frame’s orientation, yielding the rotation-compensated translational displacement $\Delta u_{\text{trans}}^{(t)}$. Because the Moiré fringes vary along one axis only, we project $\Delta u_{\text{trans}}^{(t)}$ onto the fringe-sensitive direction (perpendicular to the fringe lines) to obtain the scalar displacement signal.

We test the Moiré motion invariant (Eq. (3)) by computing the Pearson correlation between the cumulative phase signal $\{\Phi^{(t)}\}$ and the projected translational displacement $\{\Delta u_{\text{trans}}^{(t)}\}$. We evaluate the correlation over sliding windows of 30 frames, producing a per-window correlation curve across the video that serves as the authenticity signal. In the next section, we show that real and AI-generated videos produce markedly different correlation curves, suggesting ways to differentiate them.

4 Validation

We validate the Moiré motion invariant across three categories of video: real recordings, physics-based renderings, and AI-generated videos. We apply the same verification pipeline described in Sec. 3 to all categories and compare the resulting correlation signals.

4.1 Real Recorded Video

We record videos across 12 subjects in 29 distinct scenes (25 indoor, 4 outdoor) using an iPhone 15 at 1080p, 60 fps. Each recording falls into one of three motion configurations: (1) grating assembly static with camera moving, (2) grating assembly moving with camera static, and (3) both moving simultaneously (We use notation such as "Outdoor 1" to indicate an outdoor video recorded under the first condition, and so on.). The case where both are static is excluded, as our verification requires relative movement between the camera and the grating assembly to produce fringe shifts. Across subjects, we vary the movement speed and viewing distance (approximately 1.0 to 3.0 m). This yields a total of 87 videos, each approximately 15 seconds in duration.

4.2 Physics-based Rendering Video

To validate the pipeline under fully controlled conditions, we reproduce the grating assembly in Blender [3] and render videos with the Cycles physically-based rendering engine. The virtual assembly replicates the two-layer geometry described in Sec. 3, and the virtual camera matches the iPhone 15's focal length and 1080p resolution. Because Cycles traces light through the two grating layers, it accurately reproduces the parallax-driven fringe formation, providing an ideal-case for verifying the pipeline.

Each rendered sequence places the grating assembly at the world origin, sized to match the physical prototype. The camera begins at $(0.5, 0, z)$ and translates linearly to $(0, 0, z)$ over 120 frames, with small random rotations injected along the trajectory to simulate realistic camera shake. We constrain the rotations so that the grating assembly and all four ArUco markers remain fully visible in every frame. We sweep the camera distance z from 1.1 to 1.7 m in 0.1 m increments, generating 10 sequences per distance with distinct random rotation profiles, for a total of 70 rendered videos (Rendering 1 in Fig. 3). We also created a group for pure translational motion without rotation (Rendering 2 in Fig. 3). Additional experiments with different embodiments, pattern sizes, and motions involving 3D slant are provided in the supplementary material.

4.3 AI-generated Video

We evaluate AI-generated video as a potential attack against the Moiré motion invariant, testing three state-of-the-art video generation models: Veo 3.1 [16],

Grok Imagine [42] (closed-source), and LTX-2 [19] (open-source). We first attempted text-to-video (T2V) generation, providing descriptions of a person holding a grating assembly with visible Moiré fringes (replicating our real video settings). For each of the three motion configurations, we engineered specific text prompts through 15 to 20 rounds of iterative refinement. Across all models and prompt variations, none produced output containing a recognizable or trackable Moiré pattern. The generated videos either omitted the fine periodic structure or rendered it as a static, incoherent texture bearing no resemblance to real Moiré fringes, which can be identified easily by the naked eye.

We therefore focus our quantitative evaluation on the stronger image-to-video (I2V) setting, where the attacker provides one real frame of the grating assembly as conditioning input along with a text prompt describing the desired motion. We extracted the initial frame from each of the 87 authentic videos to serve as a visual condition and refined three corresponding motion prompts over 15 to 20 rounds. For each extracted frame, we use three optimized motion prompts and generate multiple outputs, yielding 321 videos in total. Because this first frame is captured from a real camera, it contains authentic Moiré fringes, giving the generator a visual prior on fringe appearance.

From the initial pool of 321 synthetic videos, manual inspection revealed that the generative models produced severe temporal inconsistencies, such as the sudden disappearance or extreme deformation of the grating assembly. We therefore screened and discarded 229 such corrupted samples, yielding a final curated dataset of 92 high-quality synthetic videos that look authentic to the human eye (see the supplementary video for examples of failed AI-generated videos).

Even in the I2V setting, most current video generation models fail to reproduce the basic visual structure of the grating assembly or the ArUco fiducial markers, and such videos would be immediately flagged as non-authentic. To stress-test the robustness of the Moiré signature itself, we adopt a deliberately generous evaluation protocol: we manually select the corner of Moiré region from the first generated frame, track it using Lucas-Kanade optical flow [4] across subsequent frames, and manually correct the tracking when it deviates. This ensures the extracted signal reflects purely the generated Moiré pattern, independent of preprocessing failures.

4.4 Results

Fig. 3 shows the distribution of Pearson correlation coefficients $|\rho|$ across all video categories. Real recordings cluster at high correlation values ($\mu = 0.87$, $\sigma = 0.14$), confirming that the Moiré motion invariant holds under diverse real-world conditions including varying lighting, distance, and motion configurations. Physics-based renderings achieve near-perfect correlations ($\mu = 0.99$), validating that the pipeline correctly captures the underlying optical relationship under ideal conditions. In contrast, AI-generated videos concentrate at substantially lower values ($\mu = 0.57$, $\sigma = 0.21$).

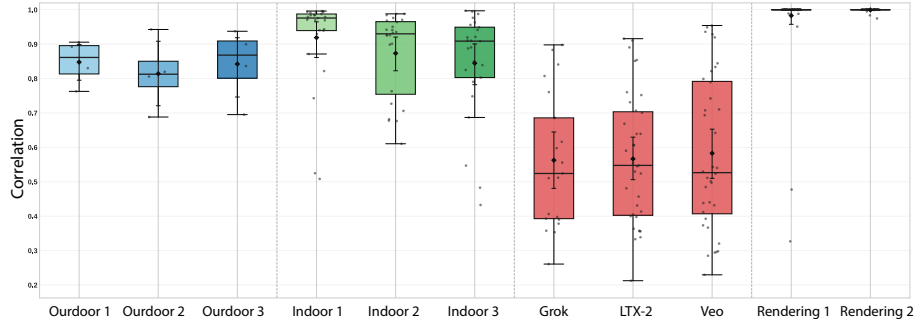


Fig. 3: Distribution of Pearson correlation coefficients $|\rho|$ across all video categories. Real recordings (indoor and outdoor) cluster at high correlation, physics-based renderings achieve near-perfect correlation, and AI-generated videos concentrate at markedly lower values.

To test whether the difference is statistically significant, we apply a Welch’s t -test, which is appropriate here because the two groups have unequal variances ($\sigma = 0.14$ vs. 0.21) and the sample sizes ($n = 86$ real, $n = 92$ AI-generated) are large enough for the Central Limit Theorem to ensure robustness to non-normality. The test confirms a highly significant separation ($t(160) = 11.6$, $p < 10^{-20}$). To quantify effect size, we compute Cohen’s $d = 1.71$, well above the 0.8 threshold conventionally considered a large effect. The gap is consistent across all three generators: Grok ($\mu = 0.56$), LTX-2 ($\mu = 0.57$), and Veo ($\mu = 0.58$) each fall far below the real-video baseline, with no model achieving a meaningful advantage.

Moreover, this results should be read as a conservative lower bound shaped by deliberate choices on both sides of the experiment. On the attack side, our protocol grants every possible advantage: image-conditioned generation with a real first frame, manual Moiré region tracking with human correction, and cherry-picked outputs representing the best case for each model. Under a less generous protocol the attacker would also need to fool ArUco marker detection and automated Moiré tracking. Also, the I2V attack is itself inherently constrained, as the attacker must first obtain a real recording of the target scene to extract a conditioning frame, significantly limiting creative freedom compared to unconstrained T2V generation. On the defense side, our prototype uses a minimal, proof-of-concept pipeline. Our tracking pipeline could be improved to reduce the influence of factors such as reflections and blur in the Moiré assembly, and to use better methods for more accurate rotation estimation. **The key take-away is therefore not the absolute threshold, but the existence of a physically grounded, measurable signal that distinguishes real from AI-generated video.**

4.5 Other Threat Models

Beyond the generative attacks evaluated above, we consider two additional threat scenarios through conceptual analysis. These threats do not require new experiments but highlight important boundaries and extensions of our approach.

Threat: Moiré region splicing. The attacker takes a clip of an authentic grating assembly from one real video and composites it into a different (potentially AI-generated) video. The fringes in the spliced region are physically correct because they originate from a real recording.

Analysis. The spliced fringes carry the phase evolution from the *source* recording, but our pipeline measures correlation against the grating assembly displacement in the *target* video. These two signals reflect different camera-assembly geometries, so they will not be linearly related. The transplanted phase signal cannot match a different motion trajectory, and the correlation will be low by construction. We omit empirical validation for this threat because the outcome is self-evident from the formulation of the invariant.

Threat: Face swap and localized editing outside the Moiré region. The attacker takes a real video of person A holding a grating assembly and applies a face-swap model (*e.g.* DeepFaceLab [27], FaceFusion [12]) to replace only the face region, making it appear as person B. The grating assembly region is left entirely untouched.

Analysis. This is the most challenging threat for our system. Because the grating assembly region is unmodified, the Moiré fringes retain their physically correct phase evolution, and the Moiré motion invariant holds. Our correlation check will report high $|\rho|$, authenticating the video. The attacker obtains a video of “person B” apparently holding person A’s grating assembly. Since the editing occurs entirely outside the Moiré region, the Moiré correlation check alone cannot detect this manipulation.

However, this class of attack is well addressed by a mature body of work on deepfake and face-swap detection [10,17,33,41]. We view these as complementary: our method authenticates whether the video was produced by a physical optical system, while editing detectors identify localized tampering. The two can be combined in a layered verification pipeline. We additionally outline two forward-looking mitigations:

Mitigation 1: Personalized Moiré identity (Moiré ID). If the grating assembly encodes a unique, verifiable identity, analogous to a public key, the identity encoded in the assembly belongs to person A, not person B. This can be achieved by fabricating the grating with specific spatial frequencies that produce a unique Moiré beat period, acting as a frequency-domain fingerprint. Alternatively, the grating can incorporate non-uniform line spacing so the fringe pattern at different spatial positions encodes distinct identifiers. The verifier then checks not only that the Moiré motion invariant holds but also that the assembly identity matches the claimed identity of the person in the video.

Mitigation 2: Spatial overlap verification protocol. In interactive settings such as video conferencing, the platform can enforce a protocol requiring

the participant to move the grating assembly across their face (*e.g.* “please wave your Moiré signature in front of your face now”). This ensures the grating assembly region spatially overlaps with the face region for at least a brief interval. Any face-swap model operating on this footage must either (a) modify the overlapping grating assembly pixels, which breaks the Moiré motion invariant, or (b) leave the assembly intact and distort the face around it, producing visible artifacts.

4.6 Discussion

Our threat analysis reveals a clear division that aligns with the current landscape of video forensics. Localized editing attacks such as face swaps fall outside the scope of what Moiré correlation alone can detect, since the grating assembly region remains untouched. However, existing deepfake and video editing detectors already achieve 82–97% accuracy on such manipulations [10, 17, 30, 33, 41], making them a problem that complementary tools can address. The harder and more urgent challenge lies in detecting fully AI-generated videos: large-scale benchmarks such as GenVidBench [30] show that cross-generator detection accuracy for T2V and I2V content often falls to 42–70%. It is precisely this gap that the Moiré motion invariant fills. No current generative model can synthesize the precise frame-to-frame coupling between fringe phase and grating assembly displacement that real optics produce, and splicing attacks are similarly defeated because the transplanted phase signal cannot match a different camera-assembly geometry. Our method thus complements existing forensic detectors in a layered verification pipeline: editing detectors screen for localized tampering, while the Moiré motion invariant provides a physics-grounded authentication layer against the generative attacks that remain beyond the reach of current detection methods.

5 Related Work

Moiré-based sensing and applications. Moiré patterns have recently been used as passive visual amplifiers for high-precision tracking. MoiréBoard [43] derives a calibration-free mapping from observed fringe geometry to 3-DoF camera position. MoiréTracker [31] characterizes screen-camera Moiré features for continuous 6-DoF camera-to-screen pose tracking. MoiréVision [32] generalizes Moiré-based sensing to curvilinear patterns under practical perspective changes. These methods use Moiré patterns as measurement signals and derive explicit motion/pose recovery models. Our work differs in both application and derivation: rather than estimating pose, we eliminate the unknown camera-motion and distance terms to obtain a correlation-based invariant between fringe phase motion and grating image motion. The resulting test requires only that the two extracted video signals remain physically coupled, making Moiré motion a passive authentication signature instead of a tracking signal.

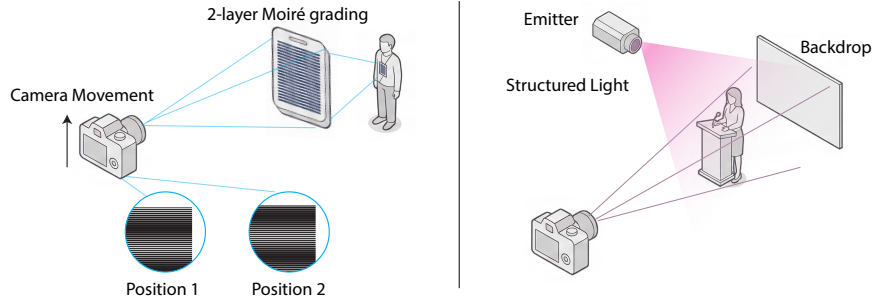


Fig. 4: (Left) Our proposed Moiré-based signature requires only a passive 2-layer Moiré structure (e.g., worn as a badge), where standard camera movement naturally captures phase shifts. (Right) Active structured light signatures [29, 35], while also using physical signals for authentication, require specialized external emitter hardware to project patterns onto the scene.

DeepFake and AI-generated video. Our work is broadly motivated by the hyper-realistic generative media that calls for reliable systems to distinguish synthetic video from real footage; DeepFake [10, 33] enables the convenient creation and dissemination of misinformation. Unlike authenticity techniques, afterthought defenses [22, 25, 26, 44, 46, 48] are in an arms race with the rapid development of generative models [5, 11, 16, 24, 28].

Physics-based video forensics. A separate line of work detects AI-generated video by identifying violations of physical plausibility, such as irregular eye pupil geometry [18], inconsistent lighting and shadow directions [13], or semantic and anatomical anomalies like extra fingers [40]. These methods exploit incidental physical errors that current generators happen to make. However, as generators improve, these artifacts are progressively eliminated through better training data and architectures. Our approach sidesteps this limitation: rather than searching for the absence of artifacts, we introduce a deliberate physical structure whose optical behavior is governed by deterministic laws that generators cannot learn to replicate statistically.

The work most closely related to ours are methods that actively project temporally modulated light at the recording site to embed a verifiable physical signature into video, such as VeriLight [35] and Michael et al. [29]. Our approach is similar in spirit, grounding authentication in physical-world signals, but differs fundamentally in that the Moiré signature is entirely passive: it arises naturally from the optical geometry of a compact grating structure and ordinary camera motion, requiring no active illumination or dedicated infrastructure (Fig. 4). A complementary line of defense involves digital watermarking [21, 45, 49], which embeds provenance information at the software level.

6 Conclusion and Limitation

Limitation. Our approach relies on the assumption that current generative models do not perform physics-based simulation of Moiré optics. Should future models incorporate full ray-tracing of the two-layer grating geometry, they could in principle reproduce the correct fringe-displacement coupling. Additionally, the method requires relative motion between the camera and the grating assembly to generate a measurable phase signal; authentication is not possible in the case where both remain static throughout the recording.

To conclude, we presented a Moiré-based video authentication method that leverages a physics-governed, deterministic link between fringe phase evolution and camera-grating motion. We derived a Moiré motion invariant that is independent of viewing distance and camera intrinsics, and built a proof-of-concept system that extracts and tests it from video alone. Experiments show significantly different correlation values for real versus AI-generated videos, enabling verifiable discrimination between authentic and synthetic recordings.

Beyond the specific Moiré-based implementation, our work points to a broader principle: physical phenomena governed by deterministic optical laws can serve as authentication signatures that are fundamentally difficult for statistical generative models to forge. We hope this perspective inspires future exploration of other physics-based signatures for media authentication, complementing existing forensic and watermarking approaches to build more robust defenses against AI-generated misinformation.

Acknowledgements

This work was supported in part by NSF 2540851 and a Gemini Academic Program Award. We thank Wenqi Wang, Yuwen Tan, Shengao Wang, Max Whitton, Michael Wakeham, Arjun Chandra, Joey Huang, Kevin Miller, and Audrey Mao for their help in testing our work in real recorded videos.

References

1. Amidror, I.: The Theory of the Moiré Phenomenon, Volume I: Periodic Layers. Springer, 2nd edn. (2009)
2. BBC: Disgraceful deep-fake AI video condemned by presidential candidate (2025)
3. Blender Foundation: Blender: Open source 3d creation suite (2025), <https://www.blender.org>
4. Bouguet, J.Y.: Pyramidal implementation of the Lucas Kanade feature tracker: Description of the algorithm. Tech. rep., Intel Corporation, Microprocessor Research Labs (2001)
5. Brooks, T., Peebles, B., Holmes, C., DePue, W., Guo, Y., Jing, L., Schnurr, D., Taylor, J., Luhman, T., Luhman, E., Ng, C., Wang, R., Ramesh, A.: Video generation models as world simulators. OpenAI Technical Report (2024)
6. ByteDance Seed Team: Seedance 2.0. ByteDance Seed (2026)

7. Campos Zamora, D., Dogan, M.D., Siu, A.F., Koh, E., Xiao, C.: MoiréWidgets: High-precision, passive tangible interfaces via moiré effect. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. pp. 1–10 (2024)
8. CNN: Finance worker pays out \$25 million after video call with deepfake chief financial officer (2024)
9. Coalition for Content Provenance and Authenticity (C2PA): C2PA technical specification. C2PA (2022)
10. Croitoru, F.A., Hiji, A.I., Hondru, V., Ristea, N.C., Irofti, P., Popescu, M., Rusu, C., Ionescu, R.T., Khan, F.S., Shah, M.: Deepfake media generation and detection in the generative AI era: A survey and outlook. arXiv preprint arXiv:2411.19537 (2024)
11. Esser, P., Chiu, J., Atighehchian, P., Granskog, J., Germanidis, A.: Structure and content-guided video synthesis with diffusion models. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 7346–7356. IEEE (2023)
12. FaceFusion: FaceFusion: Industry leading face manipulation platform. GitHub (2024)
13. Farid, H.: Lighting (in) consistency of paint by text. arXiv preprint arXiv:2207.13744 (2022)
14. Gabrielyan, E.: The basics of line moiré patterns and optical speedup. arXiv preprint arXiv:physics/0703098 (2007)
15. Google DeepMind: SynthID: Identifying AI-generated content. Google DeepMind (2023)
16. Google DeepMind: Veo 3.1 (2025), <https://deepmind.google/>
17. Gragnaniello, D., Cozzolino, D., Marra, F., Poggi, G., Verdoliva, L.: Are GAN generated images easy to detect? A critical analysis of the state-of-the-art. In: ICME (2021)
18. Guo, H., Hu, S., Wang, X., Chang, M.C., Lyu, S.: Eyes tell all: Irregular pupil shapes reveal GAN-generated faces. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 2904–2908. IEEE (2022)
19. HaCohen, Y., Brazowski, B., Chiprut, N., Bitterman, Y., Kvochko, A., Berkowitz, A., Shalem, D., Lifschitz, D., Moshe, D., Porat, E., Richardson, E., Shiran, G., Chachy, I., Chetboun, J., Finkelson, M., Kupchick, M., Zabari, N., Guetta, N., Kotler, N., Bibi, O., Gordon, O., Panet, P., Benita, R., Armon, S., Kulikov, V., Inger, Y., Shiftan, Y., Melumian, Z., Farbman, Z.: LTX-2: Efficient joint audio-visual foundation model. arXiv preprint arXiv:2601.03233 (2026)
20. Hersch, R.D., Chosson, S.: Band moiré images. ACM TOG **23**(3), 239–248 (2004)
21. Hu, R., Zhang, J., Li, Y., Li, J., Guo, Q., Qiu, H., Zhang, T.: Videoshield: Regulating diffusion-based video generation models via watermarking. In: International Conference on Learning Representations (ICLR) (2025)
22. Internò, C., Geirhos, R., Olhofer, M., Liu, S., Hammer, B., Klindt, D.: Ai-generated video detection via perceptual straightening. In: Advances in Neural Information Processing Systems. vol. 38 (2025)
23. Kafri, O., Glatt, I.: The Physics of Moiré Metrology. Wiley (1990)
24. Kuaishou Technology: Kling: A pioneering ai video generation model. <https://kling.kuaishou.com/> (2024)
25. Kundu, R., Xiong, H., Mohanty, V., Balachandran, A., Roy-Chowdhury, A.K.: Towards a universal synthetic video detector: From face or background manipulations to fully ai-generated content. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2025)

26. Li, Z., Wu, X., Shi, G., Qin, Y., Du, H., Zhou, T., Manocha, D., Boyd-Graber, J.L.: Videohallu: Evaluating and mitigating multi-modal hallucinations on synthetic video understanding. In: *Advances in Neural Information Processing Systems*. vol. 38 (2025)
27. Liu, K., Perov, I., Gao, D., Chervoniy, N., Zhou, W., Zhang, W.: Deepfacelab: Integrated, flexible and extensible face-swapping framework. *Pattern Recognition* **141**, 109628 (2023)
28. Luma AI: Dream machine: High quality, realistic video generation from text and images. <https://lumalabs.ai/dream-machine> (2024)
29. Michael, P.F., Hao, Z., Belongie, S., Davis, A.: Noise-coded illumination for forensic and photometric video analysis. *ACM Transactions on Graphics* **44**(5) (2025). <https://doi.org/10.1145/3742892>
30. Ni, Z., Yan, Q., Huang, M., Yuan, T., Tang, Y., Hu, H., Chen, X., Wang, Y.: GenVidBench: A 6-million benchmark for AI-generated video detection. In: *AAAI* (2026)
31. Ning, J., Xie, L., Li, Y., Chen, Y., Bu, Y., Wang, C., Lu, S., Ye, B.: Moirétracker: Continuous camera-to-screen 6-dof pose tracking based on moiré pattern. *IEEE Journal on Selected Areas in Communications* **42**(10), 2642–2658 (2024). <https://doi.org/10.1109/JSAC.2024.3414619>
32. Ning, J., Xie, L., Yan, Z., Bu, Y., Luo, J.: Moirévision: A generalized moiré-based mechanism for 6-dof motion sensing. In: *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. p. 467–481. *ACM MobiCom '24*, Association for Computing Machinery, New York, NY, USA (2024). <https://doi.org/10.1145/3636534.3649374>, <https://doi.org/10.1145/3636534.3649374>
33. Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., Niessner, M.: Faceforensics++: Learning to detect manipulated facial images. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (October 2019)
34. Runway: Introducing Runway Gen-4.5. Runway Research (2025)
35. Schwartz, H., Yan, X., Carver, C.J., Zhou, X.: Combating falsification of speech videos with live optical signatures. In: *Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security (CCS)*. pp. 3296–3310 (2025). <https://doi.org/10.1145/3719027.3765112>
36. Sciammarella, C.A.: The moiré method – a review. *Experimental Mechanics* **22**(11), 418–433 (1982)
37. Sethapakdi, T., Perroni-Scharf, M., Li, M., Li, J., Solomon, J., Satyanarayan, A., Mueller, S.: FabObscura: Computational design and fabrication for interactive barrier-grid animations. In: *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)* (2025)
38. Suzuki, S., Abe, K.: Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing* **30**(1), 32–46 (1985)
39. Takasaki, H.: Moiré topography. *Applied Optics* **9**(6), 1467–1472 (1970)
40. Tan, C., Ming, X., Wang, J., Tao, R., Li, B., Wei, Y., Zhao, Y., Lu, Y.: Semantic visual anomaly detection and reasoning in ai-generated images. *arXiv preprint arXiv:2510.10231* (2025)
41. Wang, S.Y., Wang, O., Zhang, R., Owens, A., Efros, A.A.: CNN-generated images are surprisingly easy to spot... for now. In: *CVPR* (2020)
42. xAI: Grok imagine video (2025), <https://x.ai>

43. Xiao, C., Zheng, C.: Moiréboard: A stable, accurate and low-cost camera tracking method. In: The 34th Annual ACM Symposium on User Interface Software and Technology. pp. 881–893 (2021)
44. Zhang, F., Li, D., Zhang, Q., Chen, J., Liu, G., Lin, J., Yan, J., Liu, J., Zha, Z.J.: Fact-R1: Towards explainable video misinformation detection with deep reasoning. In: Advances in Neural Information Processing Systems. vol. 38 (2025)
45. Zhang, K.A., Xu, L., Cuesta-Infante, A., Veeramachaneni, K.: Robust invisible video watermarking with attention. arXiv preprint arXiv:1909.01285 (2019)
46. Zhang, S., Lian, Z., Yang, J., Li, D., Pang, G., Liu, F., Han, B., Li, S., Tan, M.: Physics-driven spatiotemporal modeling for ai-generated video detection. In: Advances in Neural Information Processing Systems. vol. 38 (2025)
47. Zhang, Z.: A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence* **22**(11), 1330–1334 (2000)
48. Zheng, C., Suo, R., Lin, C., Zhao, Z., Yang, L., Liu, S., Yang, M., Wang, C., Shen, C.: D3: Training-free ai-generated video detection using second-order features. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 12852–12862 (October 2025)
49. Zou, Z., Gong, B., Wang, L.: Anti-neuron watermarking: Protecting personal data against unauthorized neural networks. In: European Conference on Computer Vision (ECCV). pp. 449–465. Springer (2022)
50. Zuiderveld, K.: Contrast limited adaptive histogram equalization. In: Graphics Gems IV, pp. 474–485. Academic Press (1994)