Nick Clark · Published April 25, 2026 · PDF

Detecting Screenshot and Recapture Fraud in Identity-Document KYC With Structural Content Identity

Remote identity verification fails when a fraudster photographs a screen, replays a leaked document image, or submits a synthetic ID that no camera ever captured, because conventional KYC pipelines score the face and read the text but cannot tell whether the pixels came from a live capture or a recaptured display. This application shows how that gap closes when a KYC onboarding flow is built on the Content Anchoring inventive step, disclosed in PCT International Application No. PCT/US26/28630, which derives a structural identity from the submitted artifact itself and reads a recapture signature directly from its gradient structure. The result is a document-authenticity check that runs at capture time, on the client, without enrolling the applicant's document image in any registry.

What This Application Specifies

Remote onboarding for banks, exchanges, lenders, and regulated marketplaces requires the applicant to submit an image of a government identity document, often alongside a selfie. Know-Your-Customer and anti-money-laundering programs, driven by requirements such as the U.S. Customer Identification Program rule under the Bank Secrecy Act and comparable obligations in other jurisdictions, treat that document image as the anchor of the applicant's claimed identity. The dominant fraud against this

step is not forging the document's printed content; it is submitting an image that was never freshly captured from a physical document at all. A fraudster photographs another person's ID displayed on a laptop screen, uploads a leaked passport scan pulled from a data breach, or feeds a generatively synthesized ID that no camera ever saw.

This application specifies a document-authenticity layer that sits inside the KYC capture flow and is built on the Content Anchoring inventive step, disclosed in PCT International Application No. PCT/US26/28630. The disclosed technology derives a unique identifier for a digital artifact from the artifact's own internal structure by extracting a multi-axis variance vector: a first axis encoding energy distribution across spatial scales, a second encoding frequency compaction, and a third, designated the Z-axis, encoding structural phase persistence based on gradient orientation distribution. Two capabilities of that disclosed pipeline carry the KYC use case. First, the screenshot recapture classifier evaluates the Z-axis gradient histogram component for the characteristic variance signature introduced when a digital display is re-captured by a camera or screen-grab, producing a recapture probability score from the candidate artifact alone. Second, the lineage query and orphan detector determine whether a submitted artifact has any registered structural lineage in a governed corpus, flagging structurally unanchored artifacts for heightened scrutiny.

No content is embedded in the applicant's document image, no enrollment of that image is required, and no central registry of citizens' documents is created. The identity is derived post-hoc from the pixels the applicant already submitted.

Why It Matters

The recapture problem is structural, and conventional KYC tooling is not equipped for it. Optical character recognition reads the document's text; face-match compares the portrait to the selfie; template checks confirm that fonts and layout match a known document design. All three can pass on a photograph of a screen showing a genuine,

stolen document, because the printed content, the face, and the layout are all authentic. What is inauthentic is the provenance of the capture, and provenance is precisely what a pixel-level content read does not surface.

The disclosed recapture detection method exploits a physical fact rather than a learned classifier. When a digital display renders an image and a camera or screen-capture device re-captures the rendered output, the display's sub-pixel geometry, the compression and dithering artifacts of the display pipeline, and the optical point-spread function of the capturing lens introduce a periodic spatial-frequency structure in the luminance channel. As disclosed, these artifacts manifest in the Z-axis gradient histogram as elevated energy in the horizontal and vertical orientation bins relative to the diagonal bins, yielding a horizontal-vertical bias score that is systematically elevated compared to the original digital artifact. The screenshot recapture classifier evaluates that bias against a policy-calibrated threshold. The disclosure states that this detection requires no reference to the original artifact and operates entirely from the structural features of the candidate itself, so no corpus lookup is needed to make the recapture call.

This matters for regulated onboarding in three concrete ways. It gives a reproducible, auditable authenticity signal rather than an opaque classifier verdict, because the disclosed admissibility decisions are replayable from the versioned policy object and the artifact's variance-derived identifier. It avoids building a database of applicants' identity images, which is itself a privacy and breach liability, because the method derives identity from the artifact rather than enrolling it. And it runs at the moment of capture rather than after the account is opened, which is where the disclosed platform positions its evaluation.

How It Composes With the Domain

A KYC onboarding flow can adopt the disclosed technology as a faithful enabling implementation as follows. The capture client is a standard web or mobile front end that already collects the document image through a browser file input or media-capture API. The disclosure specifies a client-side execution architecture in which canonical resizing, grayscale conversion, and orientation canonicalization are performed using only standard Canvas 2D interfaces, the multi-scale variance analysis and gradient-histogram computation produce a variance vector using standard arithmetic, and a 320-bit unique identifier is produced by the disclosed hash combiner, all within the browser without server-side inference or GPU compute.

At capture time the client runs the screenshot recapture classifier over the freshly computed Z-axis component. Because the recapture signal is read from the artifact itself, the raw document image need not leave the device to produce the recapture probability score; the disclosure states that the raw artifact does not leave the client during the admissibility evaluation phase and only the computed identifier and the resulting decision are transmitted. This aligns the fraud check with data-minimization expectations that constrain transmission of personal media, which is a recurring constraint in regulated identity handling.

The recapture score is one input to the disclosed composite risk score aggregator, which combines lineage absence, recapture probability, and synthesis probability into a single governance signal that routes to the pre-release admissibility engine. In the KYC framing:

- The recapture probability score flags screen-photograph and screen-grab submissions from the Z-axis horizontal-vertical bias.
- The orphan detector flags a submitted image with no registered structural lineage in the governed corpus. As disclosed, structurally unanchored artifacts are not necessarily fraudulent, but they cannot be admitted under a policy object that

requires verifiable provenance and they trigger heightened scrutiny under policy objects that govern synthetic content. For an institution that anchors its own genuine-capture reference set, an out-of-distribution submission surfaces here.

- The synthetic content detector compares the candidate variance vector against a slope-band-indexed statistical model of known synthetic-content variance profiles, producing a synthesis probability score for generatively fabricated IDs. The disclosure notes this distribution can be updated continuously as new generative architectures emerge, without retraining an inference model.

The commitment boundary is the account-approval or funding event. The disclosed pre-release admissibility engine interposes an admissibility evaluation between the candidate artifact and any irreversible or externally visible side effect, including customer delivery and downstream anchoring. A submission that trips the recapture threshold is rejected, sent back for re-capture under modified constraints, or escalated to a human reviewer through the disclosed rejection and escalation paths, each governed by a versioned, cryptographically signed policy object that defines the similarity tolerance thresholds, override authorities, and escalation routes for the relevant jurisdiction.

For institutions that resolve against a shared reference corpus, the disclosed UID resolution query protocol lets a client submit only the computed identifier, not the document, and receive an identity, derivative, orphan, or conflict resolution. Bulk resolution supports high-volume onboarding pipelines resolving many submissions per network round-trip rather than per document.

What This Enables

Adopting this layer lets a KYC program treat capture authenticity as a first-class, computable check rather than an inference left to a downstream fraud model. A screen-photograph of a stolen but genuine ID, which passes OCR, template, and face-match, is

caught by the Z-axis recapture signature that those checks never examine. A synthetic ID with no plausible capture lineage is caught by the orphan and synthesis signals feeding the composite risk score.

Because evaluation runs client-side and returns a replayable decision, an institution can demonstrate to an examiner or auditor why a given onboarding was accepted or rejected: the versioned policy object plus the artifact's variance-derived identifier reproduce the determination, in contrast to a black-box classifier score. The consultation event logger, disclosed for recording each generation event that consults a reference artifact, provides a deterministic audit record when a submission is resolved against a governed corpus, supporting the record-keeping posture that KYC and AML regimes expect. And because nothing is embedded in and no enrollment of the applicant's document is required, the institution reduces its own exposure to holding a registry of citizen identity images.

The same variance identity supports adjacent workflows without new infrastructure: detecting when the same leaked document image is replayed across many fraudulent applications through near-exact identity resolution, and detecting cropped or lightly edited derivatives of one source document through the disclosed derivative-resolution mode and per-quadrant similarity, which localizes which region of a document was altered.

Boundary Conditions

This application is faithful to what the disclosure claims and no more. The recapture classifier produces a probability score against a policy-calibrated threshold; it is a structural signal, not a legal or definitive determination of fraud, and the disclosure frames orphan and synthesis findings as triggers for heightened scrutiny rather than automatic rejection. Threshold calibration is a policy matter, and a poorly calibrated

threshold will trade false accepts against false rejects; the disclosure specifies the mechanism and the policy-object governance, not a guaranteed error rate, and no accuracy or benchmark figures are asserted here.

The recapture signal derives from display and lens artifacts of a re-captured screen. A capture pipeline that does not exhibit the disclosed elevation of horizontal-vertical gradient energy, or an adversary who deliberately reintroduces diagonal structure, may attenuate the signal; the disclosure positions recapture detection as one input to a composite risk score alongside lineage and synthesis, not as a sole determinant. Orphan detection depends on the presence and scope of a governed reference corpus; its usefulness for a given institution depends on what that institution chooses to anchor, and an unanchored artifact is expressly not equivalent to a fraudulent one. Face matching, liveness of the applicant, and document data validation are separate concerns handled by other components of a KYC stack; this layer addresses artifact-capture authenticity and structural identity, not biometric liveness.

Disclosure Scope

Every claim in this article about what the technology does traces to the Content Anchoring disclosed in PCT International Application No. PCT/US26/28630, including the multi-axis variance vector and Z-axis gradient orientation encoding, the screenshot recapture classifier and its horizontal-vertical bias signature, the orphan detector and synthetic content detector, the composite risk score aggregator, the pre-release admissibility engine and commitment boundary, the client-side execution architecture, the consultation event logger, and the UID resolution query protocol. The identity-document and KYC framing, including references to Know-Your-Customer, anti-money-laundering, the Bank Secrecy Act Customer Identification Program rule, data-minimization expectations, and remote-onboarding fraud patterns, is external domain and regulatory context provided as an enabling implementation setting; it is not part of

the disclosed invention and does not constitute legal, compliance, or regulatory advice. Named regulatory obligations are cited only as real domain context and their applicability to any particular institution is a matter for that institution's counsel.

Content Anchoring (/content-anchoring)

All 40 steps → (/inventive-steps)

Computable identity for media. Provenance from structural variance.

PCT/US26/28630 (/patents/pct-us26-28630)

PRIMARY TECHNICAL DISCLOSURE

- Content Anchoring: Computable Identity for Media That Changes (/articles/content-anchoring-computable-identity-for-media-that-changes)

SECONDARY TECHNICAL

- Multi-Axis Variance Vector Extraction: Nine Dimensions of Structural Content Identity (/articles/content-anchoring/variance-vector)
- Quadrant Decomposition: Spatial Sub-Region Fingerprinting for Partial Similarity Detection (/articles/content-anchoring/quadrant-decomposition)
- 320-Bit UID Construction: Multi-Segment Hashing for Negligible Collision Probability (/articles/content-anchoring/uid-construction)
- Structure Signature: Background-Invariant Matching Through Gradient-Only Descriptors (/articles/content-anchoring/structure-signature)
- Constellation Signature: Geometry-Invariant Matching Across Crop, Scale, and Occlusion (/articles/content-anchoring/constellation-signature)
- Five-Band Variance Classification: Content Routing by Structural Complexity (/articles/content-anchoring/variance-classification)
- Variance Saturation-Governed Cache Eviction: UID Density Replacing Static TTL (/articles/content-anchoring/cache-eviction)
- Multi-Root Composite Lineage Graphs: Provenance Through Variance Vector Similarity (/articles/content-anchoring/composite-lineage)
- Multi-Modal Content Identity: Unified Pipeline Across Image, Audio, Text, and Video (/articles/content-anchoring/multi-modal-identity)

- Rights-Grade Pre-Release Admissibility: Policy Evaluation Before Content Commitment (/articles/content-anchoring/pre-release-admissibility)
- Training Corpus Governance: Verifiable Lineage From Training Data to Model (/articles/content-anchoring/training-corpus-governance)
- Consultation Event Logging: Deterministic Records of Every Generation Reference (/articles/content-anchoring/consultation-logging)
- Model Output Provenance Fingerprint: Structural Proximity Without Model Access (/articles/content-anchoring/output-provenance)
- Creator Attribution and Compensation Routing: Payment From Consultation Lineage (/articles/content-anchoring/creator-attribution)
- Adversarial Robustness and Deepfake Detection: Content Identity as Detection Substrate (/articles/content-anchoring/adversarial-robustness)
- Client-Side Execution Architecture: Privacy-Preserving Variance Computation on Device (/articles/content-anchoring/client-side-execution)
- UID Resolution Query Protocol: Distributed Lookup Across Anchor Node Networks (/articles/content-anchoring/uid-resolution)
- Orientation Canonicalization: Rotation-Invariant Processing Through Gradient Normalization (/articles/content-anchoring/orientation-canonicalization)
- Cross-Band Resolution Pathfinding: Traversal Between Variance Bands Under Mutation (/articles/content-anchoring/cross-band-resolution)
- Identity by Position: Media as a Third Navigable Space (/articles/content-anchoring/identity-by-position)

APPLICATIONS · GENERAL

- Forbidden-Content Blocking at Upload and Generation Time: Pre-Release Exclusion Against Signed Policy (/articles/content-anchoring/forbidden-content-blocking)
- Structural Provenance for Software Supply Chains: Binary and Firmware Identity Independent of SBOM Metadata (/articles/content-anchoring/software-supply-chain-provenance)
- Rights-Grade Generative AI: How to Pay Creators, Exclude Forbidden Content, and Prevent Infringement Before Release (/articles/content-anchoring/rights-grade-generative-ai)
- Deepfake Detection by Structural Provenance: Verifying Synthetic Media Without Watermarks (/articles/content-anchoring/deepfake-provenance)
- Creator Economy Attribution Without Platform Intermediaries (/articles/content-anchoring/creator-attribution-economy)
- Verifying Source Photos and Video in the Newsroom: Content Anchoring for Journalism (/articles/content-anchoring/journalism-verification)

- Detecting Image Manipulation and Proving Figure Provenance in Research Publications (/articles/content-anchoring/academic-research-integrity)
- Content Anchoring for Legal Evidence Chains (/articles/content-anchoring/legal-evidence-chain)
- Content Anchoring for Insurance Claims Evidence (/articles/content-anchoring/insurance-claims-evidence)
- Content Anchoring for Real Estate Documentation (/articles/content-anchoring/real-estate-documentation)
- Art Authentication and Provenance Verification with Content Anchoring (/articles/content-anchoring/art-authentication)
- **Detecting Screenshot and Recapture Fraud in Identity-Document KYC With Structural Content Identity (/articles/content-anchoring/identity-document-kyc-recapture)**

APPLICATIONS · SPECIFIC

- C2PA vs Content Anchoring: Attached Provenance or Content-Intrinsic Identity? (/articles/content-anchoring/c2pa)
- Google SynthID Alternative: Content-Intrinsic Identity Beyond Watermarking (/articles/content-anchoring/google-synthid)
- Beyond Shutterstock: Content-Intrinsic Identity That Survives Re-Encoding and Cropping (/articles/content-anchoring/shutterstock)
- Spotify Alternative for Music Provenance: Structural Content Identity Beyond the ISRC Database (/articles/content-anchoring/spotify)
- Getty Images Alternative for Provenance: Structural Content Identity Beyond Metadata (/articles/content-anchoring/getty-images)
- Adobe Stock vs Structural Content Identity: Licensing Records Are Not Content Identity (/articles/content-anchoring/adobe-stock)
- YouTube Content ID vs Content Anchoring: Matching Against a Database, or Identity in the Content Itself (/articles/content-anchoring/youtube-content-id)
- Audible Magic Alternative: Structural Content Identity Beyond Database-Matched Fingerprinting (/articles/content-anchoring/audible-magic)
- Digimarc vs Structural Content Identity: Watermarks Are Added, Not Intrinsic (/articles/content-anchoring/digimarc)
- Irdeto vs Structural Content Identity: DRM Protects the Channel, Not the Payload (/articles/content-anchoring/irdeto)
- Truepic alternative: capture-time provenance versus structural identity derived from the artifact itself (/articles/content-anchoring/truepic)
- Microsoft PhotoDNA vs structural content identity: hash-matching known images versus screening artifacts before release (/articles/content-anchoring/microsoft-photodna)

- Pex alternative: structural content identity vs enrolled fingerprint matching (/articles/content-anchoring/pex)

Content Anchoring overview → (/content-anchoring)