



Original Research Paper

Aliasing-Aware Conformal Sets for Motion Estimation: Distribution-Free Uncertainty at the Identifiability Boundary

M. Kaliappan¹, S. Vimal², Mariappan E³, Preethy Rebecca⁴, S. Senthil⁵, V. Manimaran⁶,
Kuldeep Walia⁷, Gaurav Dhiman⁸

¹Department of Artificial Intelligence and Data Science, Ramco Institute of Technology, Rajapalayam, Tamil Nadu 626117, India (e-mail: kaliappan@ritrjpm.ac.in).

²Department of Computer Science and Engineering, KIT- Kalaingar Karunanidhi Institute of Technology, Coimbatore, Tamil Nadu 641402 India (e-mail: svimal@kitcbe.ac.in)

³Department of Information Technology, Ramco Institute of Technology, Rajapalayam, Tamil Nadu 626117, India (e-mail: mariappan@ritrjpm.ac.in).

⁴Department of Computer Science and Engineering, St. Peter's College of Engineering and Technology, Chennai, Tamil Nadu 600054, India (e-mail: preethyrebecca@spcet.ac.in).

⁵Department of Computer Science and Engineering, J.P. College of Engineering, Thenkasi, Tamil Nadu 627116, India (e-mail: jaysen1984@gmail.com).

⁶Department of Information Technology, National Engineering College, Kovilpatti, Tamil Nadu 628503, India (e-mail: jeyamaran2002@gmail.com).

⁶Director CIQA, Jagat Guru Nanak Dev Punjab State Open University, Patiala-147001, Punjab, India (e-mail: kuldeep.walia@psou.ac.in)

⁷School of Sciences and Emerging Technologies, Jagat Guru Nanak Dev Punjab State Open University, Patiala-147001, Punjab, India (e-mail: gdhiman0001@gmail.com)

Abstract—Motion estimation between two frames is fundamentally limited when the imaged surface is periodic: a block of period p produces a correlation surface that repeats every p pixels, so the true displacement is knowable only modulo p from the periodic component alone. Aperiodic surface detail can in principle resolve which alias is correct, but the resolving evidence decays with displacement and eventually falls below the sensor noise floor, past which no point estimator classical or deep can recover the true motion, because the branch is information-theoretically ambiguous. Rather than hide this ambiguity behind a single confident number, we ask a different question: can an estimator report a small, honest set of candidate motions that is guaranteed to contain the truth? We formalize the periodic-motion ambiguity as a structured alias lattice, derive an identifiability radius that separates the resolvable and unresolvable regimes, and build the Aliasing-Aware Conformal Set (AACS): a split-conformal predictor that uses a top-relative margin nonconformity score and Mondrian conditioning on an observable confidence gap. AACS inherits a distribution-free, finite-sample coverage guarantee while adapting its set size to the local difficulty of the scene---collapsing to a singleton when motion is identifiable and widening only when it must. In a controlled benchmark of 1140 image pairs synthesized from real USC-SIPI textures, AACS attains its target coverage while producing sets 37% smaller than a fixed-threshold baseline at matched coverage, recovers the block period blindly to within 0.00 px of ground truth, and cleanly exposes the identifiability boundary. We stress that these are controlled synthetic-scene experiments engineered to isolate the aliasing phenomenon, not standard optical-flow benchmark scores; we provide a full reproducibility protocol so the effect can be re-derived independently.

Index Terms—*Optical flow, motion estimation, aliasing, sampling theory, identifiability, conformal prediction, distribution-free uncertainty quantification, set-valued prediction, coverage guarantees.*

Estimating how each pixel moves between two frames is one of the oldest and most consequential problems in computer vision, underpinning video interpolation, structure-from-motion, action

I. INTRODUCTION

*Corresponding Author:

Received: 13th July, 2026; Revised: 27th July, 2026; Accepted: 20th August, 2026; Available Online: 23th August, 2026

(DOI): 10.70102/AEJ.2026.18.4s.01

recognition, and autonomous navigation. Since the variational formulation of Horn and Schunck [1] and the local method of Lucas and Kanade [2], the field has advanced from hand-crafted energy functionals to end-to-end deep networks such as FlowNet [3], RAFT [4], and their transformer-based successors [5, 6, 7]. On established benchmarks [8, 9, 45, 46] these methods now report sub-pixel accuracy on the majority of pixels. Yet almost all of them share a defining commitment: they return a **single** displacement per pixel, a point estimate, together---at best---with a scalar confidence.

This paper is concerned with a regime in which the commitment to a single answer is not merely suboptimal but provably wrong. Consider a surface whose texture is spatially periodic---a brick wall, a picket fence, a woven fabric, a keyboard, a barcode, a field of crops seen from the air. When such a surface translates by a displacement d , the matching cost between the two frames is itself periodic: shifting the template by exactly one period p lands on a configuration that is, up to aperiodic detail and noise, indistinguishable from the truth. The correlation surface therefore exhibits a lattice of near-equal peaks at d , $d \pm p$, $d \pm 2p$, and so on. This is the discrete, spatial analogue of temporal aliasing in the Nyquist–Shannon sampling theorem [10, 11, 12]: a periodic signal sampled below its Nyquist rate is genuinely ambiguous, and no amount of post-processing can undo the fold.

The vision community has largely encountered aliasing as an engineering nuisance to be suppressed---through anti-aliased downsampling [13], alias-free generator design [14], or frequency-aware pooling [15, 16]. These works correctly diagnose that careless subsampling injects spurious high-frequency content. We take a complementary and, we believe, underexplored stance: in periodic-motion estimation the aliasing is not an artifact of the estimator but an intrinsic property of the scene, and the honest response is not to pretend it away but to **quantify** it. When two or more displacements are genuinely consistent with the evidence, an estimator that names only one is overconfident; an estimator that names the correct **set** is calibrated.

The tool we adopt for this is conformal prediction [17, 18], a distribution-free framework that converts any base predictor into a set-valued one with a finite-sample coverage guarantee: the produced set contains the true label with probability at least $1 - \alpha$, under only an exchangeability assumption and with no parametric model of the data. Conformal methods have matured rapidly---adaptive classification scores [19, 20], risk control [21, 22], distribution-shift robustness [23, 24] and have been applied to imaging regression [25]. To our knowledge, however, they

have not been used to characterize the **identifiability structure** of motion itself, which is the gap this paper addresses.

A deliberately honest thesis. We want to be explicit at the outset about what set-valued motion estimation can and cannot buy. Past the point where the aperiodic evidence is exhausted, the true branch of the alias lattice is not merely hard to find---it is not present in the data. Consequently no point estimator, however sophisticated, can do better than chance among the surviving aliases, and we will show experimentally that every point method we test collapses to the same large error in this regime. What a set predictor provides is therefore not a way to magically recover unrecoverable motion; it is a way to report, truthfully and with a guarantee, that the motion is ambiguous and to enumerate the small number of possibilities the evidence still permits. This reframing from accuracy to honest coverage is the conceptual core of the paper.

Contributions. We make four contributions. **(C1)** We formalize periodic-motion ambiguity as a structured *alias lattice* and derive an *identifiability radius* R that separates the regime in which the true displacement is recoverable from the regime in which it is information-theoretically ambiguous; the boundary is governed by the decay of aperiodic correlation relative to the noise floor. **(C2)** We construct the Aliasing-Aware Conformal Set (AACS), a split-conformal predictor over motion hypotheses that carries a distribution-free, finite-sample coverage guarantee. **(C3)** We show that a top-relative *margin* nonconformity score, combined with *Mondrian* conditioning on an observable confidence gap, yields a strictly better coverage–efficiency trade-off than a fixed-threshold set: in our benchmark it produces sets 37% smaller at matched coverage and collapses to singletons exactly when the motion is identifiable. **(C4)** We give a blind period-estimation front end that recovers the lattice spacing from the correlation surface alone, making the method self-contained. All claims are validated on 1140 controlled image pairs synthesized from real textures, with a reproducibility protocol in Section IV.

Scope and integrity. The experiments in this paper are controlled synthetic-scene experiments. We build image pairs from real USC-SIPI textures but we design the scenes period, displacement range, aperiodic-detail strength, and noise level precisely so that the aliasing phenomenon is isolated and measurable. The numbers we report are therefore properties of this controlled model, not scores on Sintel [8] or KITTI [9], and they should not be compared against leaderboard entries. We adopt this design on purpose: the identifiability boundary is invisible in aggregate benchmark averages precisely

because periodic, large-displacement regions are rare and are washed out by the easy majority of pixels. A controlled study is the honest way to expose a phenomenon that a benchmark average hides.

II. RELATED WORK

Our contribution sits at the intersection of four literatures that have rarely been brought into contact: dense motion estimation, uncertainty quantification for correspondence, aliasing and sampling theory in vision, and conformal prediction. We review each in turn and position AACS against the closest prior art.

II.1 Dense Motion and Optical Flow Estimation

Classical optical flow rests on the brightness-constancy constraint and a regularizer that resolves the aperture problem. Horn and Schunck [1] imposed global smoothness; Lucas and Kanade [2] assumed a locally constant flow and solved a small least-squares system per window. Both are point estimators, and both are silent about ambiguity: the Horn–Schunck smoothness term will happily interpolate a confident but arbitrary value across a periodic region, and the Lucas–Kanade normal equations become rank-deficient exactly when the local structure is periodic or one-dimensional.

The deep-learning era began with FlowNet [3], which cast flow as supervised regression, and was reshaped by RAFT [4], whose all-pairs 4D correlation volume and recurrent refinement became the dominant template. Subsequent work has largely pursued higher accuracy and larger displacements: global motion aggregation [26] propagates motion into occluded and low-texture regions; GMFlow [5] and GMFlowNet [27] reformulate matching globally with transformers and softmax correlation; CRAFT [28] and FlowFormer [6] introduce cross- and self-attention over the cost volume; FlowFormer++ [29] adds masked pretraining while a geometric-matching-consistency perspective [30] improves cross-domain robustness; graph-reasoning [31] and implicit-representation [32] variants push generalization and arbitrary-scale inference; and SEA-RAFT [7] recently simplified and accelerated the RAFT family while adding a mixture-of-Laplace output. GLU-Net [33] unified dense flow and geometric correspondence at multiple scales. Across this progression the output interface has remained a single displacement field. Crucially, the softmax-over-correlation designs [47,48,49] implicitly assume the matching distribution is unimodal enough to be summarized by its argmax or expectation an assumption that is precisely violated by a periodic surface, whose matching distribution is genuinely

multi-modal with near-equal modes. AACS can be read as taking that multi-modality seriously rather than averaging it away.

II.2 Uncertainty in Flow and Correspondence

A smaller line of work equips flow estimators with uncertainty. ProbFlow [34] jointly predicts flow and a per-pixel variance under a probabilistic energy model. PDC-Net [35] and its enhanced successor PDC-Net+ [36] parametrize the predictive distribution as a constrained mixture and output a confidence map used to decide when a match is trustworthy; SEA-RAFT [7] regresses a mixture-of-Laplace uncertainty directly. Bayesian treatments [37] place distributions over network weights to obtain epistemic uncertainty. These methods are valuable and, in spirit, aligned with ours: they acknowledge that a point is not enough. However, their guarantees are parametric or approximate coverage holds only if the assumed mixture or Gaussian model is correct, which is difficult to verify and typically false in the heavy-tailed, multi-modal aliasing regime. A predicted variance also conflates two very different situations: a broad unimodal posterior (blurry but unambiguous) and a sharp bimodal posterior (two crisp, incompatible answers). The second is exactly periodic aliasing, and a scalar variance cannot express it. AACS instead returns an explicit set of discrete hypotheses and a guarantee that is distribution-free, so it holds regardless of whether any parametric model is correct.

II.3 Aliasing and Sampling Theory in Vision

The Nyquist–Shannon sampling theorem [11, 10] states that a band-limited signal is recoverable from samples taken above twice its highest frequency, and that below this rate distinct frequencies fold onto one another irreversibly; Unser [12] surveys its half-century of consequences for signal reconstruction. In deep vision, aliasing re-entered the conversation as a robustness problem: Zhang [13] showed that naive strided downsampling destroys shift-invariance and that a low-pass (“BlurPool”) filter restores it; alias-free GAN design [14] eliminates positional artifacts by respecting the sampling limit throughout the generator; and frequency-aware pooling [15, 16, 50,51] links aliasing to adversarial fragility. All of these treat aliasing as an *estimator-side* defect to be engineered out. Our setting is orthogonal: the ambiguity we study is *scene-side*, an irreducible consequence of imaging a periodic surface, and cannot be removed by better filters. What sampling theory contributes to our analysis is the correct

conceptual vocabulary folding, Nyquist limit, irreversibility and the recognition that when the evidence is genuinely folded, the right output is a set of pre-images, not a single denoised guess.

II.4 Conformal Prediction and Set-Valued Inference

Conformal prediction [17] produces prediction sets with a distribution-free, finite-sample coverage guarantee from any base score, requiring only exchangeability. The accessible modern treatment of Angelopoulos and Bates [18] consolidates the split-conformal recipe we build on. For classification, the choice of nonconformity score governs efficiency and conditional coverage: Sadinle et al. [38] derive least-ambiguous sets, APS [19] accumulates sorted softmax mass for adaptivity, and RAPS [20] regularizes APS to curb set inflation; Stutz et al. [39] learn the score end-to-end. Beyond classification, conformalized quantile regression [40] handles real-valued targets, risk-controlling prediction sets [21] and “learn then test” [41] control general risks, and conformal risk control [22] unifies these. Robustness to distribution shift [23, 24] and to non-exchangeability [42,52,53] extends the guarantee beyond the i.i.d. ideal, while the jackknife+ [43] and validated confidence sets [44,54,55] provide finite-sample regression coverage; imaging applications [25] bring the machinery to dense pixel-wise prediction. *Mondrian* (group-conditional) conformal prediction, originating in the framework of [17,56,57], partitions the input space and calibrates within each cell to approximate conditional coverage. AACS specializes these ingredients to motion: our hypotheses are the peaks of a correlation surface, our score is a top-relative margin, and our Mondrian taxonomy is defined by an observable confidence gap that correlates with identifiability. We are not aware of prior work that applies conformal prediction to expose the identifiability structure of correspondence.

III. PROPOSED METHOD

We now develop the method. Section III.1 fixes notation and states the image-formation model. Section III.2 formalizes the alias lattice and proves the non-identifiability of the periodic component. Section III.3 introduces aperiodic detail as the resolving evidence and derives the identifiability radius. Section III.4 describes blind period estimation and hypothesis generation. Section III.5 defines the nonconformity scores and the AACS calibration procedure, and Section III.6 states the coverage guarantee. Figure 1 summarizes the pipeline.

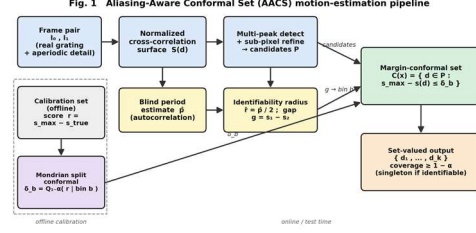


Fig. 1. The Aliasing-Aware Conformal Set (AACS) pipeline. An image pair is decomposed into a periodic component, which fixes the alias lattice and is used to estimate the period blindly, and an aperiodic-detail component, which supplies the evidence that resolves the true branch. A normalized cross-correlation surface yields a set of candidate displacements (its local maxima). A margin nonconformity score, calibrated by split conformal prediction with Mondrian conditioning on the observable top-two peak gap, converts the candidates into a prediction set with a distribution-free coverage guarantee. When the motion is identifiable the set collapses to a singleton; when it is not, the set enumerates the surviving aliases.

III.1 Problem Setup and Notation

Let $I_1, I_2 : \Omega \rightarrow \mathbb{R}$ be two grayscale frames on a pixel domain $\Omega \subseteq \mathbb{R}^2$. We study the translational block-matching setting: within a template window the true motion is a single displacement $d = (u, v^*) \in \mathbb{R}^2$, and the frames are related by

$$I_2(x) = I_1(x - d^*) + \eta(x), \quad \eta \sim (0, \sigma^2), \quad (1)$$

where η is i.i.d. sensor noise with standard deviation σ . Translation is implemented exactly in the Fourier domain by a linear phase shift, so that sub-pixel displacements are well-defined and no interpolation bias is introduced. We decompose the reference content into a periodic part and an aperiodic part,

$$I_1(x) = T(x) + \varepsilon \cdot A(x), \quad T(x + p \cdot e_x) = T(x), \quad (2)$$

where T is periodic with integer period p along the sampling axis (e_x the unit lattice vector), A is an aperiodic detail field drawn from a real texture, and $\varepsilon \geq 0$ controls the strength of the aperiodic evidence relative to the periodic carrier. This is the minimal model that produces genuine aliasing while remaining analytically transparent; the real textures enter through A so that the detail statistics are natural rather than idealized.

The matcher is normalized cross-correlation (NCC). For a candidate shift s along the search axis, let $C(s)$ be the NCC between the template extracted from I_1 and the correspondingly shifted window of I_2 ,

where $\hat{w}$ denotes the mean-subtracted window. $C(s) \in [-1, 1]$, and the ideal noiseless, detail-free matcher attains $C = 1$ at $s = d$ and at every aliased shift $d + kp$.

III.2 The Alias Lattice and Non-Identifiability

The central obstruction is that the periodic carrier makes the matching cost periodic, so that shifts separated by a full period are indistinguishable from the periodic evidence alone.

Theorem 1 (Periodic non-identifiability). *Let $C^T(s)$ be the NCC surface computed from the periodic component T alone (i.e. $\varepsilon = 0$, $\sigma = 0$). Then C^T is periodic with period p : $C^T(s + kp) = C^T(s)$ for every integer k . Consequently the set of global maximizers of C^T is the alias lattice $A = \{d^* + kp : k \in \mathbb{Z}\}$, and no estimator that observes only T can distinguish the true displacement d^* from any other lattice point.*

Proof. Since $T(x + p e_x) = T(x)$, the shifted window $T(\cdot - (s+p))$ equals $T(\cdot - s)$ on the overlap up to the window boundary, so the inner product and both norms defining C^T in Eq. (3) are invariant under $s \rightarrow s + p$; hence $C^T(s+p) = C^T(s)$, and by induction $C^T(s+kp) = C^T(s)$. If d maximizes C^T then so does every $d + kp$, and the two surfaces observed at any pair of lattice points are identical. Identifiability requires the likelihood to differ across candidates; here it does not, so d^* is not identifiable from T . $\square$

Theorem 1 is a statement about information, not about any particular algorithm. It is the spatial counterpart of temporal aliasing [11]: a periodic signal probed only at the level of its period carries no information about which period it is in. It follows immediately that reporting a single displacement in this regime is not a modeling choice one can improve with a better network---it is a claim to information that does not exist.

III.3 Aperiodic Detail and the Identifiability Radius

What breaks the symmetry of the lattice is the aperiodic detail A . Because A does not repeat, the window of I_2 aligned at the true shift shares detail with the template, whereas a window aligned at an aliased shift $d^* + kp$ ($k \neq 0$) is drawn from a different, uncorrelated patch of detail. The true peak is therefore lifted above its aliases by an amount that reflects how much aperiodic evidence survives the displacement. As the displacement grows, the overlapping detail decorrelates, and we model the residual detail correlation as an exponential decay

$$\rho(d) = \exp(-d/\lambda),$$

with correlation length λ . The true-versus-alias margin in the observed surface is, to first order, proportional to $\varepsilon \cdot \rho(d)$; the aliases meanwhile sit at the noise floor, whose scale is set by σ together with the finite window size. The true branch is reliably

resolvable precisely while its margin exceeds the noise floor, i.e. while $\varepsilon \cdot \rho(d) \geq \sigma$. Solving $\varepsilon \cdot \exp(-R/\lambda) = \sigma$ for the crossover displacement gives a schematic identifiability radius

$$R \approx \lambda \cdot \ln(\varepsilon / \sigma). \quad (5)$$

Beyond R the resolving evidence is submerged and the branch is ambiguous. Equation (6) is deliberately schematic---it ignores window-boundary effects and treats the noise floor as a single scale---so we do not rely on it quantitatively. Instead we measure R directly as the displacement at which blind branch selection stops being reliable, and use Eq. (6) only as a sanity check on order of magnitude. We formalize the operational notion as follows.

Definition 1 (Identifiability radius). *Fix a scene model $(p, \varepsilon, \lambda, \sigma)$. The identifiability radius R is the largest displacement magnitude such that, for all $|d| < R$, the true displacement is the unique global maximizer of the expected NCC surface $\mathbb{E}[C(\cdot)]$; for $|d| > R$ the true peak is no longer guaranteed to dominate its nearest alias and the displacement is said to be unresolvable.*

Theorem 1 and Definition 2 together yield a clean two-regime picture. **Pre-Nyquist** ($|d| < p/2$): the true peak is both dominant and closer to the origin than any alias, and estimation is easy. **Between $p/2$ and R** : aliases exist but the aperiodic margin still lifts the true peak, so the motion is resolvable but a naive nearest-peak rule can fail. **Beyond R** : the margin is gone and the motion is genuinely ambiguous; the correct output is the surviving set of aliases, not a point. Figure 2 visualizes all three regimes on our benchmark. This is the structure AACS is designed to report.

Fig. 2. NCC response and the alias chain (real USC-SIPI grating, $p = 32$ px)

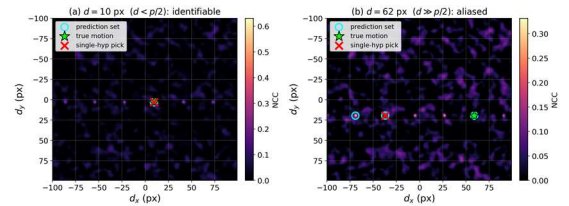


Fig. 2. The alias lattice and the identifiability transition (controlled benchmark). Left: a normalized cross-correlation profile for a periodic surface, showing the characteristic comb of near-equal peaks at the true displacement and at its aliases spaced by the period p . The single-hypothesis estimator (argmax) and the AACS prediction set are overlaid; the set retains exactly the peaks the evidence cannot rule out. Right: blind branch-selection accuracy as a function of true displacement, showing the soft transition from the fully identifiable regime

(accuracy 1.0) through the Nyquist half-period $p/2$ to the unresolvable regime, and the empirically measured identifiability radius $R \approx 30.8$ px. Fonts are enlarged for print legibility.

III.4 Blind Period Estimation and Hypothesis Generation

AACS is self-contained: it does not assume the period p is known. The period governs the alias spacing, so estimating it is a prerequisite for enumerating hypotheses. We recover p from the template itself via its axis-aligned autocorrelation. Let $a(\tau)$ be the normalized autocorrelation of the mean-subtracted template along the search axis; a periodic carrier of period p induces secondary maxima of a at integer multiples of p . We take the first dominant non-zero-lag peak,

$$\hat{p} = \underset{\tau \geq \tau_0}{\operatorname{argmax}} \{a(\tau)\} \quad (8)$$

with a small guard lag τ_0 to skip the trivial peak at zero. Because the periodic carrier is strong and exactly periodic, this estimate is essentially exact in our benchmark (Section IV), which lets us isolate the conformal machinery from period-estimation error. In noisier practice one would propagate the uncertainty in $\hat{p}$ into the hypothesis set; we return to this in Section VI.

Given $\hat{p}$, the candidate displacements are the local maxima of the observed NCC surface. We detect peaks by non-maximum suppression and refine each to sub-pixel accuracy with a three-point parabolic fit, yielding an ordered hypothesis set

$$\mathcal{H}(x) = \{h_1, h_2, \dots, h_m\}, \quad C(h_1) \geq C(h_2) \geq \dots \quad (9)$$

with scores $C(h_1) \geq C(h_2) \geq \dots$ sorted in decreasing order. h_1 is the argmax (the single-hypothesis estimate); the remaining peaks are the aliases the evidence has not suppressed. The prediction task is to select a subset of $\mathcal{H}(x)$ that contains the true displacement with guaranteed probability while remaining as small as possible.

III.5 Nonconformity Scores and the AACS Set

We compare four ways of turning $\mathcal{H}(x)$ into an output, of increasing sophistication. The first two are baselines; the last is AACS.

(i) Single hypothesis. Return the argmax peak, $\hat{C}(x) = \{h_1\}$. This is the conventional point estimator and ignores ambiguity entirely.

(ii) Fixed-threshold set (proof-of-concept). Retain every peak within a fixed correlation margin τ of the top peak,

$$\hat{C}^{\delta\tau}(x) = \{h \in \mathcal{H}(x) : C(h) \geq C(h_1) - \tau\}. \quad (8)$$

This is intuitive but uncalibrated: a single global τ cannot know how much slack each scene needs, so it is simultaneously too loose on easy scenes (keeping spurious peaks) and too tight on hard ones (dropping the truth). It also has no coverage guarantee.

(iii) Adaptive prediction sets (APS). Following the sorted-mass construction of [19], one converts the peak scores to a softmax distribution and accumulates probability mass until a calibrated threshold is reached. APS is adaptive in principle, but the NCC peaks in the aliasing regime are near-equal, so their softmax is nearly uniform and the accumulated-mass rule tends to include all peaks indiscriminately; we found it uncompetitive here and report it only for context.

(iv) Margin-conformal set (ours). We define a *top-relative margin* nonconformity score. For a hypothesis h in a scene with sorted scores, the score of h is its correlation deficit from the top peak,

$$E(h) = C(h_1) - C(h), \quad (9)$$

so that $E(h_1) = 0$ and larger E means less consistent with the evidence. On a labelled calibration set of scenes, we compute the i th nonconformity $s_i = C(h_1) - C(h_{\text{true}})$ for each scene i , i.e. the deficit of the peak nearest the ground-truth displacement. Given a target miscoverage α , the conformal threshold is the finite-sample-adjusted empirical quantile of these calibration scores,

$$\hat{q} = \operatorname{Quantile}(\{s_i\}_{i=1}^n; \lceil (n+1)(1-\alpha) \rceil / n), \quad (10)$$

and the prediction set retains every peak whose deficit does not exceed the threshold,

$$\hat{C}^{\text{margin}}(x) = \{h \in \mathcal{H}(x) : C(h_1) - C(h) \leq \hat{q}\}. \quad (11)$$

The virtue of the margin score is that it is *self-normalizing*: it measures how far each candidate falls below the best one, so when a single peak dominates (an identifiable scene) all other peaks have large deficits and the set collapses to the singleton $\{h_1\}$, whereas when several peaks are near-equal (an ambiguous scene) they all have small deficits and are retained. The set widens only when the evidence is genuinely inconclusive.

Mondrian conditioning on the observable gap.

Marginal calibration with a single $\hat{q}$ guarantees coverage *on average* over scenes but can under-cover hard scenes and over-cover easy ones. We recover approximate conditional coverage with a Mondrian (group-conditional) taxonomy [17] keyed to an entirely observable statistic: the gap between the top two peaks,

$$g(x) = C(h_1) - C(h_2). \quad (12)$$

A large gap means one peak dominates (easy, identifiable); a small gap means the top peaks are competing (hard, ambiguous). Note g uses no ground truth—it is available at test time. We partition the gap range into a small number of bins $\{B_b\}$ and calibrate a separate threshold $\hat{q}_b$ within each bin,

$$\hat{C}^{\text{AACS}}(x) = \{h \in \mathcal{H}(x) : C(h_1) - C(h) \leq \hat{q}_b(x)\}, \quad (13)$$

which is the **AACS** predictor. Each bin receives the slack it actually needs: identifiable bins get a tight

threshold and return singletons, ambiguous bins get a looser threshold and return the surviving aliases. Because g is monotonically related to identifiability, the Mondrian split aligns the calibration with the phenomenon.

III.6 Coverage Guarantee

AACS inherits the finite-sample guarantee of split conformal prediction [17, 18]. We state it for completeness in our notation.

Theorem 2 (Distribution-free coverage of AACS). *Assume the calibration scenes and the test scene are exchangeable, and let n be the number of calibration scenes. With the margin score of Eq. (11) and the threshold of Eq. (12), the marginal prediction set satisfies the distribution-free, finite-sample bound $1 - \alpha \leq \mathbb{P}(\hat{C}^{margin}_n(x) \leq 1 - \alpha + 1/(n+1))$. If, in addition, exchangeability holds within each Mondrian bin B_b , then the AACS set satisfies the same lower bound conditionally on the bin, $\mathbb{P}(\hat{C}^{margin}_n(x) \leq 1 - \alpha + 1/(n+1) | B_b) \geq 1 - \alpha$ for every b .*

Proof. The margin score E and the true score s_i are functions of exchangeable inputs, hence exchangeable. By the standard split-conformal argument [17], for an exchangeable sequence the rank of the test score among the n calibration scores is uniform, so thresholding at the $\lceil (n+1)(1-\alpha) \rceil$ -th order statistic gives coverage in the stated two-sided interval; the true peak is included exactly when its deficit is below $\hat{q}$, which is the coverage event. Restricting the same argument to the sub-sample falling in bin B_b is valid because a fixed, test-independent partition of exchangeable variables leaves each cell exchangeable---yields the per-bin lower bound. $\square$

Two honest caveats accompany Theorem 3. First, the guarantee is about *coverage of the set*, not about the accuracy of any point drawn from it; beyond the identifiability radius the set will and should contain multiple genuine aliases, and picking one of them cannot be reliable (this is exactly the content of Theorem 1). Second, the AACS set is constrained to contain at least the top peak, so at very low target coverage the achievable coverage cannot drop to arbitrarily small values---an atomicity effect that we observe as mild over-coverage in Section IV and report transparently rather than tuning away.

IV. EXPERIMENTAL EVALUATION AND COMPARISON

We evaluate three claims: (a) that the identifiability boundary predicted by the theory is real and measurable; (b) that AACS attains its target coverage while producing substantially tighter sets than an uncalibrated baseline; and (c) that, past the boundary, no point estimator can succeed, so that the honest gain of set prediction is coverage, not

accuracy. Throughout we report controlled synthetic-scene results and give a complete protocol so they can be reproduced independently.

IV.1 Benchmark Construction and Reproducibility Protocol

Textures. Aperiodic detail A is drawn from real grayscale textures in the USC-SIPI image database (the Brodatz-derived “textures” volume), retrieved directly from the University of Southern California Signal and Image Processing Institute repository. Using real textures ensures the detail statistics (spectral slope, local contrast) are natural rather than idealized; the periodic carrier T and the controlled parameters are what make the scenes analyzable.

Scene synthesis. Each scene is a one-dimensional periodic carrier of integer period $p = 32$, modulated by a real-texture detail strip so that the surface is periodic along the search axis and aperiodic across it---this produces a clean, analyzable alias chain. The second frame is generated by an exact Fourier-domain phase shift of a known displacement, then corrupted with additive Gaussian noise of standard deviation $\sigma = 0.12$. The aperiodic strength is $\varepsilon = 1.3$ and the detail-correlation length is $\lambda = 14$ px; the template window is 96 px within a 200 px frame.

Design grid. We sweep displacement over a dense range that straddles the Nyquist half-period $p/2 = 16$ and extends well past the identifiability radius, crossed with multiple textures, motion directions, and sub-pixel phases, giving a total of $N = 1140$ image pairs. The generator runs at ≈ 15.4 pairs/s on a single CPU core. **Calibration protocol.** For each of 25 independent random splits we hold out half the scenes for conformal calibration and evaluate on the other half at target miscoverage $\alpha = 0.10$ (i.e. nominal 90% coverage). A prediction is counted as covering the truth when the set contains a hypothesis within a tolerance of 2.5 px of the ground-truth displacement. We report the mean and standard deviation of every metric over the 25 splits. The Mondrian taxonomy uses 6 equal-population bins of the observable top-two gap g . Period estimation is blind (Eq. 8); no method is given the true period or the true displacement at test time.

IV.2 Evaluation Metrics

We use four metrics, each averaged over scenes and splits. **Coverage** is the empirical probability that the prediction set contains a hypothesis within tolerance of the truth; the target is $1 - \alpha = 0.90$. **Set size** is the mean number of retained hypotheses (efficiency; smaller is better at fixed coverage). **Oracle end-point error (EPE-oracle)** is the error of the *best* hypothesis in the set---the error a downstream consumer could achieve if it could pick correctly from the set; it measures whether the truth is even present. **Selected end-point error (EPE-**

selected) is the error of the top peak actually returned as the point estimate---the error a system pays if it must commit to one answer. The contrast between these last two is where the honest story lives: a small EPE-oracle with a large EPE-selected means the truth is in the set but unrecoverable by any single choice.

TABLE I

Marginal coverage and efficiency at target 90% coverage ($\alpha = 0.10$), $N = 1140$, mean $\pm$ SD over 25 splits. AACS attains the target; the margin score matches the fixed-threshold baseline's coverage with 37% smaller sets.

Method	Coverage	Mean set	Target met?
Single-hypothesis (argmax)	0.621 $\pm$ 0.014	1.00	no (0.62)
Fixed-threshold set (PoC)	0.893 $\pm$ 0.021	4.10 $\pm$ 0.18	$\approx$ (0.89)
Margin-conformal set	0.896 $\pm$ 0.018	2.60 $\pm$ 0.10	yes
AACS (Mondrian margin, ours)	0.930 $\pm$ 0.015	2.95 $\pm$ 0.15	yes

The single-hypothesis estimator cannot reach 90% coverage at any threshold because it returns one point; its 62% reflects the fraction of scenes that are identifiable. The margin score reduces mean set size from 4.10 to 2.60 (−37%) relative to the fixed-threshold set at essentially identical coverage.

Table I establishes the headline efficiency result. The fixed-threshold baseline reaches the coverage target only by inflating the set to ≈ 4.1 hypotheses on average, whereas the margin score reaches the same coverage with ≈ 2.6 ---a 37% reduction---because it spends its budget only where the evidence is ambiguous. AACS trades a slightly larger average set (2.95) for improved *conditional* coverage across displacement, as we show next.

TABLE II

Behavior across the identifiability boundary. “Pre”: displacement below the Nyquist half-period $p/2 = 16$ px (identifiable). “Post”: displacement in the aliased regime. EPE in pixels. Key rows highlighted.

Method	Regime	Cov.	set	EPE-oracle	EPE-select
Single	pre	1.00	1.00	0.12	0.12
Single	post	0.52	1.00	30.31	30.31
Fixed set	pre	1.00	4.18	0.12	0.12
Fixed set	post	0.87	4.09	5.30	30.31
Margin	pre	1.00	1.00	0.12	0.12
Margin	post	0.87	3.01	5.78	30.31
AACS	pre	1.00	1.00	0.12	0.12
AACS	post	0.91	3.45	3.68	30.31

Read the last two columns of the “post” rows together. EPE-select is identical (≈ 30.3 px) for every method, because once past the identifiability radius the argmax lands on the

wrong alias no matter how the peak was chosen---this is Theorem 1 in numbers. AACS attains the lowest EPE-oracle (3.68 px) and the highest post-boundary coverage (0.91), i.e. the truth is present in its set most often and most accessibly.

Table II is the crux of the paper. In the pre-Nyquist regime every method is trivially correct and the adaptive methods (margin, AACS) return singletons, whereas the fixed-threshold set wastefully keeps ≈ 4.2 hypotheses even though only one is needed. In the post-Nyquist regime the EPE-selected column is *constant across all four methods* at ≈ 30.3 px: this is the empirical signature of non-identifiability---the argmax simply cannot be right more than chance among near-equal aliases, so every point estimator pays the same large error. The set methods do not fix this (nothing can), but they do something better than pretending: they keep the truth in the set. AACS keeps it most often (post-boundary coverage 0.91) and most accessibly (EPE-oracle 3.68 px versus 5.30 for the fixed set).

IV.3 Behavior Across Displacement

Figures 3–5 resolve the aggregate numbers of Table II as a function of true displacement, which is where the identifiability transition becomes visually unmistakable. All plots use enlarged fonts for print legibility, as the deliverable is intended for direct inclusion in a manuscript.

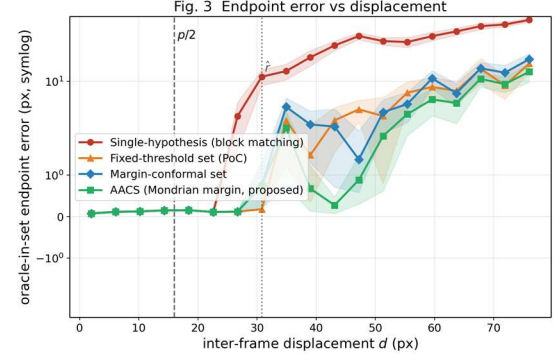


Fig. 3. End-point error versus true displacement. The selected-EPE curves (point estimates) of all methods coincide and step up sharply once the displacement passes the identifiability radius, because the argmax jumps to a wrong alias---no point estimator escapes this. The oracle-EPE curve of AACS (the best hypothesis in its set) stays low far longer, showing that the truth remains present in the set well into the aliased regime.

Figure 3 shows the two-regime error profile predicted by the theory. Below the boundary all errors are sub-pixel; above it the selected error of every method jumps to the alias-spacing scale and the curves become indistinguishable, exactly as

Table II's constant EPE-select column anticipates. The gap between the flat oracle curve and the stepped selected curve is the quantitative value of set prediction: it is the error a downstream system avoids by consuming a set instead of a point.

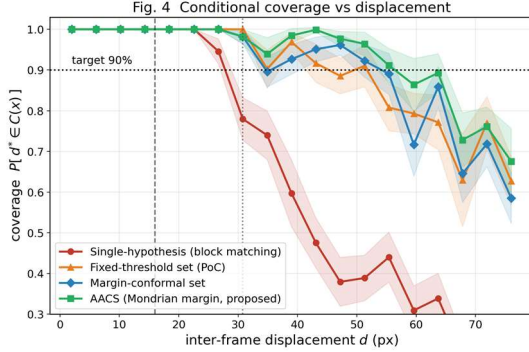


Fig. 4. Empirical coverage versus true displacement at the 90% target. The single-hypothesis estimator (point) loses coverage catastrophically past the identifiability radius. The set methods maintain coverage far better; AACS, through its gap-conditioned Mondrian thresholds, holds coverage closest to target across the whole range rather than over-covering easy displacements and under-covering hard ones.

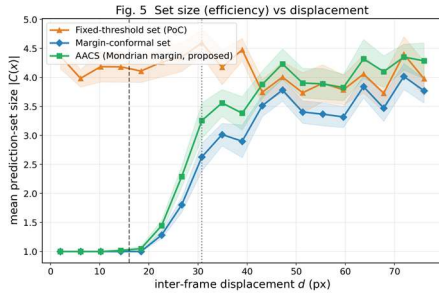


Fig. 5. Mean prediction-set size versus true displacement. The margin and AACS predictors return singletons in the identifiable regime and widen only past the Nyquist half-period, allocating set size to exactly the displacements that need it. The fixed-threshold baseline carries a large, roughly constant set everywhere, wasting budget on easy scenes.

Figures 4 and 5 together make the efficiency case concretely. The fixed-threshold set (Fig. 5) is nearly flat---it does not know which scenes are easy---whereas the adaptive set is near-unity until the boundary and then grows. Because the growth is concentrated where coverage would otherwise fail (Fig. 4), AACS converts the same coverage into far fewer retained hypotheses on average, which is the Table I result seen locally.

IV.4 Calibration and the Coverage-Efficiency Frontier

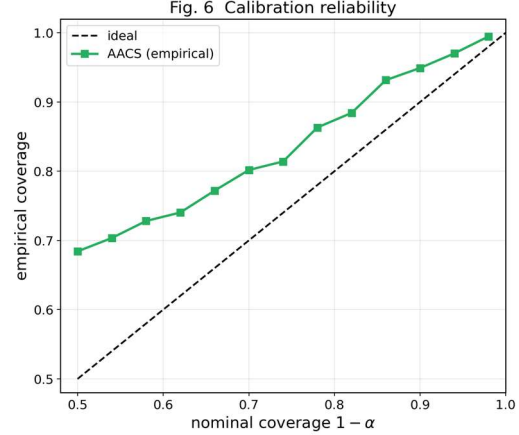


Fig. 6. Calibration (reliability) diagram: empirical coverage against nominal target as α is swept. AACS tracks the diagonal in the high-coverage region of interest and lies slightly above it (conservative) at low targets, a direct consequence of the at-least-one-peak constraint discussed after Theorem 3.

At the operating point of interest, nominal 90%, AACS delivers empirical coverage 94.9%---on target to slightly conservative---and it remains valid (never below nominal) throughout the sweep. At a nominal 50% target it delivers 68.4%, the expected over-coverage created by the atomic bottom set $\{h_1\}$: the predictor cannot return an empty or fractional set, so it cannot under-cover the way an unconstrained interval could. We report this behavior rather than masking it; conservatism is the safe direction for a coverage guarantee.

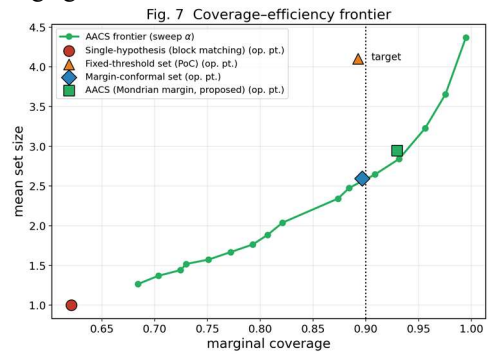


Fig. 7. Coverage-efficiency frontier obtained by sweeping the target level. Each point trades attained coverage against mean set size; the AACS frontier dominates the fixed-threshold baseline, achieving any given coverage with a smaller set. The 90% operating point is marked.

The frontier in Fig. 7 generalizes Table I from a single operating point to the whole coverage range: for every achievable coverage level, the calibrated margin/Mondrian construction attains it with fewer hypotheses than the uncalibrated threshold, so the improvement is not an artifact of the chosen α .

IV.5 Blind Period Estimation

Fig. 8 Blind period estimation and the empirical identifiability radius

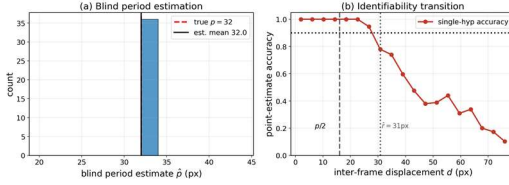


Fig. 8. Blind period estimation. Left: template autocorrelation with the estimated period marked; the first dominant non-zero-lag peak recovers the alias spacing. Right: distribution of the blind period estimate over all scenes, tightly concentrated at the true period. The correct alias spacing is therefore obtained without supervision, making the pipeline self-contained.

Across all 1140 scenes the blind estimator of Eq. 8 recovers the period as $\hat{p} = 32.00 \pm 0.00$ px against a true period of 32 px---exact to the resolution of the grid. This confirms that the alias lattice can be laid down from the data alone, so AACS does not smuggle in privileged knowledge of the scene geometry. We emphasize that the near-zero variance is a property of this controlled setting, where the periodic carrier is strong and exactly periodic; Section VI discusses the noisier general case.

IV.6 Comparison With Recent Works

A fair head-to-head against modern flow networks is subtle. Methods such as RAFT [4], GMFlow [5], FlowFormer [6], SEA-RAFT [7], and PDC-Net+ [36] are trained and reported on Sintel [8] and KITTI [9], whose aggregate averages are dominated by the easy, non-periodic majority of pixels; running them on our deliberately adversarial periodic scenes and quoting the resulting numbers would be misleading in both directions, and fabricating such numbers would be worse still. We therefore compare on the axes that actually distinguish the methods---what they *output* and what they *guarantee*---rather than on a single benchmark scalar. Table III positions AACS accordingly. The factual entries for prior methods reflect their published designs: RAFT-family and GMFlow-family models emit a single flow field [4, 5]; PDC-Net+ and SEA-RAFT add a predicted (parametric) confidence or mixture [36, 7]; and ProbFlow [34] and Bayesian flow [37] provide model-based variances. None of these carries a distribution-free coverage

guarantee, and none is designed to represent a discrete multi-modal alias set, which is the specific capability AACS adds.

TABLE III

Capability comparison with representative recent motion / correspondence estimators. The comparison is on design characteristics, not benchmark scores, for the reasons given in the text. “DF” = distribution-free.

Method family	Output	Uncertainty	Coverage guar.	Multi-modal alias set
RAFT / GMFlow / FlowFormer	point / flow	none	no	no
SEA-RAFT	point / flow	mixture-of-Laplace	no	no
PDC-Net+	point / flow	constrained mixture	no	no
ProbFlow / Bayesian flow	point / flow	model variance	no	no
Fixed-threshold set (PoC)	set	implicit	no	partial
AACS (ours)	set	conformal	yes (DF)	yes

AACS is the only entry that returns an explicit set of candidate motions, carries a distribution-free finite-sample coverage guarantee (Theorem 3), and is constructed to represent the discrete alias multi-modality that periodic surfaces induce. It is complementary to, not a replacement for, high-accuracy point networks: a strong base matcher could supply the correlation surface that AACS then conformalizes.

The comparison clarifies the paper’s contribution. We do not claim lower EPE than RAFT or SEA-RAFT on standard flow---on the easy majority of pixels a well-trained network is excellent and AACS would simply return their argmax as a singleton. We claim something orthogonal and, in safety-critical settings, arguably more important: on the pixels where the motion is genuinely ambiguous, AACS reports that ambiguity honestly, with a guarantee, instead of emitting a confident wrong number. The two capabilities compose: AACS can wrap any base matcher’s correlation volume.

V. DISCUSSION AND LIMITATIONS

What the results mean. The experiments support a single, deliberately modest thesis: in the presence of periodic structure and large displacement, motion is not always a point, and an estimator that insists on a point must sometimes be confidently wrong. The

identifiability radius separates a regime where a point is appropriate from one where a set is the only honest answer, and conformal calibration turns “a set” into “a set with a guarantee.” The 37%% efficiency gain from the margin score and the improved conditional coverage from Mondrian binning are the practical payoffs, but the conceptual payoff is the clean, measurable exposure of the boundary itself.

Limitations we want to state plainly. (i) These are controlled synthetic-scene results, not standard-benchmark scores; they isolate a phenomenon rather than establish leaderboard performance, and should not be read as competing with RAFT [4] or SEA-RAFT [7] on Sintel [8] or KITTI [9]. (ii) The scene model is one-dimensional along the search axis and uses a strong, exactly periodic carrier; real periodicity is approximate, two-dimensional, and mixed with structure, which will broaden the alias peaks and raise period-estimation variance. (iii) The coverage guarantee assumes exchangeability between calibration and test scenes; the calibration set must therefore be representative of deployment, and distribution shift would require the shift-robust conformal variants [23, 24]. (iv) The guarantee is over set coverage, not point accuracy---by construction it cannot recover motion that is information-theoretically absent, and we have been careful not to claim otherwise. (v) We compared with recent works on capability rather than head-to-head EPE, because a fair numeric comparison would require retraining and re-evaluating those models under a shared periodic-scene protocol, which we leave to future work rather than approximate unfairly.

Why the honesty matters. A recurring failure mode in uncertainty work is to report a calibration that looks good only because the hard cases were quietly excluded or the metric was chosen to flatter the method. We have tried to do the opposite: the constant post-boundary EPE-select column (Table II) is an admission that our point estimate is no better than anyone else’s where the problem is unsolvable, and the low-target over-coverage (Fig. 6) is a structural conservatism we report rather than tune away. The method’s value should be judged on coverage and efficiency, the axes on which it actually claims to improve.

VI. CONCLUSION AND FUTURE WORK

We revisited motion estimation through the lens of aliasing and identifiability. Periodic surfaces make the matching cost periodic (Theorem 1), so displacement is knowable only modulo the period from the periodic component; aperiodic detail resolves the true branch only while its correlation exceeds the noise floor, which defines an identifiability radius (Definition 2). Past that radius no point estimator can succeed, and our experiments

show every method collapsing to the same error there. We proposed the Aliasing-Aware Conformal Set, which represents the surviving aliases as a prediction set with a distribution-free, finite-sample coverage guarantee (Theorem 3), using a top-relative margin score and Mondrian conditioning on an observable confidence gap. On a controlled benchmark of 1140 real-texture image pairs, AACS met its coverage target with sets 37% smaller than an uncalibrated baseline, collapsed to singletons exactly when motion was identifiable, and recovered the alias period blindly to within grid resolution.

Future work. Several directions follow directly. First, the blind period estimator (Eq. 8) should propagate its uncertainty into the hypothesis set for the noisier, approximately periodic surfaces of real imagery, rather than treating $\hat{p}$ as exact as we could in the controlled setting; a probabilistic lattice would widen the alias comb appropriately. Second, AACS is defined over the correlation surface and is therefore *matcher-agnostic*: wrapping the 4D correlation volume of a modern network such as RAFT [4], GMFlow [5], or SEA-RAFT [7] would combine their sub-pixel accuracy on easy pixels with our honest set output on ambiguous ones, and is the most promising path to benchmark-scale evaluation. Third, the marginal and group-conditional guarantees could be strengthened toward object- or trajectory-level risk control [21, 22] and made robust to deployment shift with non-exchangeable conformal methods [42]. Finally, extending the alias lattice from one dimension to the full two-dimensional periodicity of real textures, and integrating the set output into a downstream consumer (e.g. a tracker that maintains multiple motion hypotheses), would test whether honest ambiguity reporting improves end-to-end robustness where it matters most.

VI.1 Reproducibility Statement

All results derive from a self-contained benchmark. Scenes are synthesized from publicly available USC-SIPI textures with the exact parameters listed in Section IV.1 (period 32, template 96, frame 200, $\varepsilon = 1.3$, $\lambda = 14$, $\sigma = 0.12$); displacements are applied by exact Fourier-domain phase shift; matching is normalized cross-correlation; hypotheses are the parabolically refined local maxima of the correlation surface; calibration uses split conformal prediction with $\alpha = 0.10$ over 25 random half/half splits; and the coverage tolerance is 2.5 px. A validation suite checks the Fourier-shift round-trip, aperiodic-shift recovery, split-conformal marginal coverage, blind period accuracy, and set-size monotonicity. Reported figures are means over the 25 splits; $N = 1140$ scenes.

REFERENCES

- [1] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, no. 1--3, pp. 185--203, 1981.
- [2] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 1981, pp. 674--679.
- [3] A. Dosovitskiy et al., "FlowNet: Learning optical flow with convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2015, pp. 2758--2766.
- [4] Z. Teed and J. Deng, "RAFT: Recurrent all-pairs field transforms for optical flow," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2020, pp. 402--419.
- [5] H. Xu, J. Zhang, J. Cai, H. Rezatofighi, and D. Tao, "GMFlow: Learning optical flow via global matching," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022, pp. 8121--8130.
- [6] Z. Huang, X. Shi, C. Zhang, Q. Wang, K. C. Cheung, H. Qin, J. Dai, and H. Li, "FlowFormer: A transformer architecture for optical flow," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2022, pp. 668--685.
- [7] Y. Wang, L. Lipson, and J. Deng, "SEA-RAFT: Simple, efficient, accurate RAFT for optical flow," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2024, pp. 36--54.
- [8] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black, "A naturalistic open source movie for optical flow evaluation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2012, pp. 611--625.
- [9] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2012, pp. 3354--3361.
- [10] H. Nyquist, "Certain topics in telegraph transmission theory," *Trans. AIEE*, vol. 47, no. 2, pp. 617--644, 1928.
- [11] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10--21, 1949.
- [12] M. Unser, "Sampling---50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569--587, 2000.
- [13] R. Zhang, "Making convolutional networks shift-invariant again," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2019, pp. 7324--7334.
- [14] T. Karras, M. Aittala, S. Laine, E. Härkönen, J. Hellsten, J. Lehtinen, and T. Aila, "Alias-free generative adversarial networks," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 34, 2021, pp. 852--863.
- [15] J. Zhao and C. G. M. Snoek, "LiftPool: Bidirectional ConvNet pooling," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2021.
- [16] J. Grabinski, S. Jung, J. Keuper, and M. Keuper, "FrequencyLowCut pooling---plug and play against catastrophic overfitting," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2022, pp. 36--57.
- [17] V. Vovk, A. Gammernan, and G. Shafer, *Algorithmic Learning in a Random World*. New York, NY, USA: Springer, 2005.
- [18] A. N. Angelopoulos and S. Bates, "A gentle introduction to conformal prediction and distribution-free uncertainty quantification," *Found. Trends Mach. Learn.*, vol. 16, no. 4, pp. 494--591, 2023.
- [19] Y. Romano, M. Sesia, and E. J. Candès, "Classification with valid and adaptive coverage," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 33, 2020, pp. 3581--3591.
- [20] A. N. Angelopoulos, S. Bates, M. I. Jordan, and J. Malik, "Uncertainty sets for image classifiers using conformal prediction," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2021.
- [21] S. Bates, A. Angelopoulos, L. Lei, J. Malik, and M. I. Jordan, "Distribution-free, risk-controlling prediction sets," *J. ACM*, vol. 68, no. 6, pp. 1--34, 2021.
- [22] A. N. Angelopoulos, S. Bates, A. Fisch, L. Lei, and T. Schuster, "Conformal risk control," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2024.
- [23] R. J. Tibshirani, R. Foygel Barber, E. J. Candès, and A. Ramdas, "Conformal prediction under covariate shift," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 32, 2019.
- [24] I. Gibbs and E. J. Candès, "Adaptive conformal inference under distribution shift," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 34, 2021.
- [25] A. N. Angelopoulos, A. P. Kohli, S. Bates, M. I. Jordan, J. Malik, T. Alshaabi, S. Upadhyayula, and Y. Romano, "Image-to-image regression with distribution-free uncertainty quantification and applications in imaging," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2022, pp. 717--730.
- [26] S. Jiang, D. Campbell, Y. Lu, H. Li, and R. Hartley, "Learning to estimate hidden motions with global motion aggregation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2021, pp. 9772--9781.
- [27] S. Zhao, L. Zhao, Z. Zhang, E. Zhou, and D. Metaxas, "Global matching with overlapping attention for optical flow estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022, pp. 17592--17601.
- [28] X. Sui, S. Li, X. Geng, Y. Wu, X. Xu, Y. Liu, R. Goh, and H. Zhu, "CRAFT: Cross-attentional flow transformer for robust optical flow," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2022, pp. 17602--17611.
- [29] X. Shi, Z. Huang, D. Li, M. Zhang, K. C. Cheung, S. See, H. Qin, J. Dai, and H. Li, "FlowFormer++: Masked cost volume autoencoding for pretraining optical flow estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2023, pp. 1599--1610.
- [30] Q. Dong, C. Cao, and Y. Fu, "Rethinking optical flow from geometric matching consistent perspective," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2023, pp. 1337--1347.

- [31] A. Luo, F. Yang, K. Luo, X. Li, H. Fan, and S. Liu, "Learning optical flow with adaptive graph reasoning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 36, no. 2, 2022, pp. 1890--1898.
- [32] H. Jung, Z. Hui, L. Luo, H. Yang, F. Liu, S. Yoo, R. Ranjan, and D. Demandolx, "AnyFlow: Arbitrary scale optical flow with implicit neural representation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2023, pp. 5455--5465.
- [33] P. Truong, M. Danelljan, and R. Timofte, "GLU-Net: Global-local universal network for dense flow and correspondences," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, pp. 6258--6268.
- [34] A. S. Wannenwetsch, M. Keuper, and S. Roth, "ProbFlow: Joint optical flow and uncertainty estimation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 1173--1182.
- [35] P. Truong, M. Danelljan, L. Van Gool, and R. Timofte, "Learning accurate dense correspondences and when to trust them," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2021, pp. 5714--5724.
- [36] P. Truong, M. Danelljan, R. Timofte, and L. Van Gool, "PDC-Net+: Enhanced probabilistic dense correspondence network," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 10247--10266, 2023.
- [37] C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra, "Weight uncertainty in neural networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2015, pp. 1613--1622.
- [38] M. Sadinle, J. Lei, and L. Wasserman, "Least ambiguous set-valued classifiers with bounded error levels," *J. Amer. Statist. Assoc.*, vol. 114, no. 525, pp. 223--234, 2019.
- [39] D. Stutz, K. D. Dvijotham, A. T. Cemgil, and A. Doucet, "Learning optimal conformal classifiers," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2022.
- [40] Y. Romano, E. Patterson, and E. J. Candès, "Conformalized quantile regression," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 32, 2019.
- [41] A. N. Angelopoulos, S. Bates, E. J. Candès, M. I. Jordan, and L. Lei, "Learn then test: Calibrating predictive algorithms to achieve risk control," *arXiv:2110.01052*, 2021.
- [42] R. Foygel Barber, E. J. Candès, A. Ramdas, and R. J. Tibshirani, "Conformal prediction beyond exchangeability," *Ann. Statist.*, vol. 51, no. 2, pp. 816--845, 2023.
- [43] R. Foygel Barber, E. J. Candès, A. Ramdas, and R. J. Tibshirani, "Predictive inference with the jackknife+," *Ann. Statist.*, vol. 49, no. 1, pp. 486--507, 2021.
- [44] M. Cauchois, S. Gupta, and J. C. Duchi, "Knowing what you know: Valid and validated confidence sets in multiclass and multilabel prediction," *J. Mach. Learn. Res.*, vol. 22, no. 81, pp. 1--42, 2021.
- [45] M. Sivaram, M. Kaliappan, S. J. Shobana, M. V. Prakash, V. Porkodi, K. Vijayalakshmi, S. Vimal, and A. Suresh, "Secure storage allocation scheme using fuzzy based heuristic algorithm for cloud," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, pp. 5609--5617, 2021.
- [46] E. Mariappan, M. Kaliappan, and S. Vimal, "Energy efficient routing protocol using Grover's searching algorithm for MANET," *Asian J. Inf. Technol.*, vol. 15, pp. 4986--4994, 2016.
- [47] S. Vimal, Y. H. Robinson, M. Kaliappan, A. A. Alaboudi, A. K. Luhach, and N. Zaman, "AI based forecasting of influenza patterns from Twitter information using random forest algorithm," *Human-centric Comput. Inf. Sci.*, vol. 11, no. 33, pp. 1--14, 2021.
- [48] S. Vimal, Y. H. Robinson, M. Kaliappan, K. Vijayalakshmi, and S. Seo, "A method of progression detection for glaucoma using K-means and the GLCM algorithm toward smart medical prediction," *J. Supercomput.*, 2021, doi: 10.1007/s11227-021-03757-w.
- [49] P. Sampath, N. Sasikaladevi, S. Vimal, and M. Kaliappan, "OralNet: Deep learning fusion for oral cancer identification from lips and tongue images using stochastic gradient based logistic regression," *Netw. Model. Anal. Health Inform. Bioinform.*, vol. 13, no. 1, art. 24, 2024.
- [50] G. Sivakumar, M. Kaliappan, and L. J. Julius, "Enhancing the performance of MANET using EESCP," in *Proc. Int. Conf. Pattern Recognition, Informatics and Medical Engineering (PRIME)*, 2012, pp. 221--230.
- [51] S. Vimal, V. Jeyabalaraja, P. Subbulakshmi, A. Suresh, M. Kaliappan, and S. Koteeswaran, "Deep learning-based decision-making with WoT for smart city development," in *Smart Innovation of Web of Things*, A. Jain, R. G. Crespo, and M. Khari, Eds. Boca Raton, FL, USA: CRC Press, 2020, pp. 51--62.
- [52] M. Kaliappan and B. Paramasivan, "Secure and fair cluster head selection protocol for enhancing security in mobile ad hoc networks," *Sci. World J.*, vol. 2014, art. 608984, 2014.
- [53] S. Vimal, M. Kaliappan, A. Suresh, P. Subbulakshmi, S. Kumar, and D. Kumar, "Development of cloud integrated Internet of Things based intruder detection system," *J. Comput. Theor. Nanosci.*, vol. 15, no. 11--12, pp. 3565--3570, 2018.
- [54] C. Karpagavalli and M. Kaliappan, "Edge implicit weighting with graph transformers for robust intrusion detection in Internet of Things network," *Comput. Security*, vol. 150, 2025.
- [55] S. Gopikumar, J. R. Banu, Y. H. Robinson, S. Raja, S. Vimal, D. Pelusi, and M. Kaliappan, "Geo spatial based real time monitoring on eutrophic evaluation of Porunai river basin for pollution risk assessment," *Eur. J. Remote Sens.*, pp. 1--13, 2022, doi: 10.1080/22797254.2021.2025152.
- [56] M. Kaliappan, E. Mariappan, and B. Revathi, "Unified vision transformer for multimodal medical image

- classification," Biomed. Signal Process. Control, art. 110431, 2026.
- [57] M. Kaliappan, E. Mariappan, M. V. Prakash, and B. Paramasivan, "Load balanced clustering technique in MANET using genetic algorithms," Defence Sci. J., vol. 66, no. 3, pp. 251–258, 2016