

Evaluating the Stability of Spatial Keypoints via Cluster Core Correspondence Index

Suvadip Mukherjee, Thibault Lagache, Jean-Christophe Olivo-Marin

▶ **To cite this version:**

Suvadip Mukherjee, Thibault Lagache, Jean-Christophe Olivo-Marin. Evaluating the Stability of Spatial Keypoints via Cluster Core Correspondence Index. IEEE Transactions on Image Processing, 2020, 30, pp.386-401. 10.1109/TIP.2020.3036759 . pasteur-03699899

HAL Id: pasteur-03699899

https://pasteur.hal.science/pasteur-03699899

Submitted on 20 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Evaluating the Stability of Spatial Keypoints via Cluster Core Correspondence Index

Suvadip Mukherjee, *Member, IEEE*, Thibault Lagache and Jean-Christophe Olivo-Marin, *Fellow, IEEE*

***Abstract*—Detection and analysis of informative keypoints is a fundamental problem in image analysis and computer vision. Keypoint detectors are omnipresent in visual automation tasks, and recent years have witnessed a significant surge in the number of such techniques. Evaluating the quality of keypoint detectors remains a challenging task owing to the inherent ambiguity over what constitutes a good keypoint. In this context, we introduce a reference based keypoint quality index which is based on the theory of spatial pattern analysis. Unlike traditional correspondence-based quality evaluation which counts the number of feature matches within a specified neighborhood, we present a rigorous mathematical framework to compute the statistical correspondence of the detections inside a set of salient zones (cluster cores) defined by the spatial distribution of a reference set of keypoints. We leverage the versatility of the level sets to handle hypersurfaces of arbitrary geometry, and develop a mathematical framework to estimate the model parameters analytically to reflect the robustness of a feature detection algorithm. Extensive experimental studies involving several keypoint detectors tested under different imaging scenarios demonstrate efficacy of our method to evaluate keypoint quality for generic applications in computer vision and image analysis.**

***Index Terms*—spatial analysis, keypoint detection, density estimation, level set.**

I. Introduction

COMPUTER vision based automation tasks such as image retrieval, object tracking, homography estimation and scene reconstruction often build on efficient detection and representation of local image features [1]. A fundamental prerequisite in many such applications is therefore the reduction of high-dimensional digital signal into a parsimonious set of distinguishable and informative local measurements or *keypoints*. This paper discusses a fundamental problem of evaluating keypoint quality using a quantitative measure based on the theory of spatial pattern analysis.

Efficient encoding of the underlying image semantics require the extracted keypoints to be distinguishable and sufficiently invariant to perturbations [2]. Typical applications involving keypoint detectors include homography estimation between scenes [3] via interest point matching [4], robust tracking of objects in digital video [5]–[7] and feature based image retrieval [8] from digital databases. In addition to such traditional computer vision tasks, efficient estimation of local keypoints is necessary for applications involving unstructured data such as 3D point clouds [9], [10]. Typically, a keypoint detection system needs to address two critical subproblems. First, a detector is used to localize feature points of interest which are then associated with a local descriptor to encapsulate key information either using supervision or via unsupervised feature modeling. [7], [11]–[14]. Naturally, efficacy of the system relies heavily on the quality of the extracted keypoints. This enforces a strict criteria to ensure robustness of keypoint detectors for such diverse and multi-domain applications, therefore motivating an objective and methodical assessment of their quality.

Quantitative assessment is critical for systematic *technology evaluation* [15] because it presents an unbiased methodological approach to measure the efficacy of a solution. For example, denoising techniques in image analysis are commonly evaluated via the structural similarity metric [16], while segmentation accuracy can be quantified using the Dice coefficient or the Jaccard similarity index [17]. However, there exists few techniques to quantitatively analyze the performance of a keypoint detection system. Objective evaluation of keypoint detectors is considered a hard problem, because keypoint localizations routinely correspond to abstract and complex semantics (such as blobs, corners) which are not reliably identified by the human visual system. Therefore, in the absence of gold-standard manual annotations, the generic strategy is to evaluate keypoints' robustness via a *stability* index [4]. The fundamental question here is *"how do we define a mathematically meaningful notion of stability for keypoint detectors?"* Intuitively, stability is a measure of robustness, and robust detectors produce keypoints which are largely immune to imaging artifacts and perturbations such as noise, image contrast and geometric distortions [3], [12]. Common evaluation metrics define stability as a measure of correspondence between a reference set of keypoints (derived from an image under normal conditions) versus a *perturbed* set of detections which are either obtained when the image is subject to various modes of imaging artifacts (such as additive noise, illumination changes, etc.) or related by homography [3], [4]. For robust detectors, a strong measure of association is expected between these two sets of spatial data, while a lack of association implies instability.

The main objective of this research is to develop a reference-based mathematical index of stability for generic keypoint detectors, independent of the application. We address a particular drawback of traditional correspondence based measures which typically count the number of matched keypoints within a predefined local neighborhood. Such counting based approaches [4], [18] do not reflect stability in the true sense, since positive correspondence may be observed even when the keypoints are randomly distributed. This over-estimation is particularly pronounced for dense keypoint localizations, and

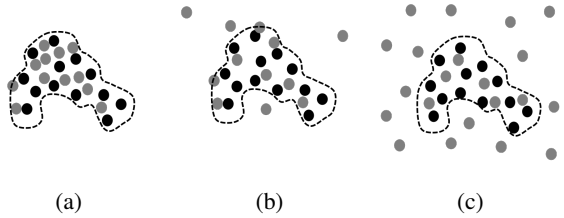
The authors are with the Bioimage Analysis Unit, Institut Pasteur, 25, rue du Docteur Roux, 75015 Paris, France. Email: suvadip.mukherjee@pasteur.fr

has been termed as 'quantity bias' in [19].We define a notion of statistical correspondence of the perturbed keypoints to a set of salient zones (namely cluster cores) derived from the reference distribution in an unsupervised, null hypothesis framework. The proposed **C**luster **C**ore **C**orrespondence **I**ndex or C^{3}I is designed to eliminate bias accumulating through false positive correspondences due to randomness in keypoint localizations. Based on the theory of spatial data analysis [20], *C^{3}I quantifies detector stability as a function of significant accumulation of keypoints to the cluster cores which is statistically above the expected value when the keypoint distribution follows a homogeneous Poisson process (null hypothesis) [21].* Compared to the probabilistic methods for quantifying keypoint deviation [22], [23], we find that C^{3}I is both robust to small imperceptible changes in feature localizations, while being sensitive to more global changes. Another feature of this methodology is that the proposed is computed analytically without resorting to computationally demanding Monte-Carlo techniques for parameter estimation.

C^{3}I is fundamentally a robust alternative to traditional measures of keypoint stability. Therefore, it would serve as a generic tool to diagnose, and quantitatively evaluate the quality of any keypoint detector, irrespective of the specific downstream application. To the best of our knowledge, this is the first work to present an objective keypoint evaluation methodology based on the principle of statistical analysis of the spatial detections. In the next section common challenges in quantifying keypoint stability are discussed, followed by a brief review of relevant research techniques. Technical details of the proposed work are detailed in section III. We also observe positive correlation between keypoint stability via C^{3}I and the end-product quality, which supports our hypothesis that stable keypoints are essential for robust automation tasks. Furthermore, C^{3}I could also be viewed as a statistical affinity measure between spatial point sets which could potentially be useful for applications beyond keypoint stability measurement. A few such relevant examples, along with elaborate experimental analyses are furnished in section IV.

II. Background and Motivation

High level semantic concepts such as image features and keypoints are difficult to formalize in an application independent fashion. Keypoints are often informally defined as *"local, invariant geometric structures in images such as blobs or corners"* [1], although this could vary depending on the specific application espcially when keypoints are defined in the context of unstructured data [9], [10]. Such keypoints are typically extracted either via efficient geometrical modeling of the image manifold [12], [13], [27], [28], or by supervised machine learning techniques [14], [29]–[31]. To ensure a robust end product, Thacker et al. [15] advocate a set of best practices for performance evaluation of computer vision systems. The authors emphasize on rigorous quantitative validation of the techniques under different conditions such as changing illumination, noise levels and imaging artifacts. Mokhtarian and Mohanna [32] define a keypoint consistency index that penalizes change in absolute number of keypoint detections, along with an accuracy measure. However, this evaluation is restricted only to corner detectors because it is impossible to reach a consensus on precise keypoint locations by human judges for generic feature detectors. Traditional reference based evaluation methods determine the level of association between two sets of spatial keypoints $P = \{\mathbf{p}\}$ and $Q = \{\mathbf{q}\}$, $\mathbf{p}, \mathbf{q} \in \mathbb{R}^2$. The set P is typically considered to be the *reference*, which correspond to the keypoint detector's response to an unperturbed imaging system. The *perturbed* point set Q represents the detector's response to external artifacts such as noise, illumination change, rotation etc. Therefore, for an ideal system it is expected that $P \approx Q$, where set similarity is measured according to some spatial deviation index. The pioneering works of Schmid et al. [33] and Mikolajczyk et al. [4], and the corner detector performance measure in [18] propose a measure of local correspondence between spatial datasets which are related by known homography. In addition to such techniques, other probabilistic measures of correspondence have also been developed which are based on spatial analysis of the keypoint distributions [23], [34]. In the following subsection, we present an overview of some relevant reference based keypoint stability evaluators.



(a) (b) (c)

Fig. 1: Three hypothetical keypoint detection results are presented which depict increasing levels of instability from (a)-(c). For each image, the reference set of points P are shown in black, and the gray points correspond to perturbed detections Q. The cluster core contours are shown as the black dotted contour. An ideal detector is presented in (a) where the perturbed points are completely embedded within the cluster cores. The detector in (b) performs slightly worse than the one in (a) since a few localizations are outside the cluster cores. Finally, (c) illustrates an unstable system, where keypoints localizations are random.

A. Local correspondence based keypoint stability analysis

In this category of evaluation methods, the stability index between $P = \{\mathbf{p}\}$ and $Q = \{\mathbf{q}\}$ is proportional to the total number of one-to-one correspondences between the two sets. A prominent work to evaluate keypoint detectors is due to Schmid *et al.* [33] which identifies the fraction of keypoints which are repeated in the common regions of two images. To account for the uncertainty in detection, the authors propose a stability index ρ_s to detect correspondence in a local r-neighborhood (where r is a user-defined local distance parameter) in the following manner:

$$\rho_s(r) = |\mathcal{L}(r)| / \min\{|P|, |Q|\} \quad (1)$$
$$\text{where, } \mathcal{L}(r) = \{(\mathbf{p}, \mathbf{q}) : \left|\left|\mathbf{p} - \mathbf{q}\right|\right| \leq r\}$$

Despite its popularity, finding such local correspondence is complicated by dense accumulation of the keypoints, and makes this index sensitive to the neighborhood size. Another overlap-based stability criteria is used in [18] which is conceptually similar to the Mander's overlap coefficient [35]. This

is written as $\rho_m(r) = |P_r \cap Q_r| / \min\{|P_r|, |Q_r|\}$, where $P_r = \{\cup \mathrm{p}_r\}$ where $\mathrm{p}_r = \{\mathbf{x} \in \mathbb{R}^2 : ||\mathbf{x} - \mathbf{p}||_2 \leq r\}$ and Q_r is also defined similarly. Again, the parameter r is difficult to estimate, although there have been a few rule-of-thumb proposals [4]. Lack of precision in parameter selection impacts the evaluation of the detection system. For example, significantly small value of r would result in elevated false negatives and would impact specificity, while the number of false positives increases if r is too large. Other feature matching criteria such as the nearest-neighbor approach used in [36] also rely on ad-hoc threshold parameters which are not easy to estimate. Also, the aforementioned measures do not take into account the potential stochasticity of detector algorithms, and such robustness measures are biased due to the false positive correspondence of the perturbed keypoints.

B. Keypoint stability via spatial density analysis

In contrast to local keypoint correspondence analysis, a few works have adopted a framework to identify the change in detections by comparing the probability density functions derived from the spatial positions. Zhang *et al.* [37] and Jaremo *et al.* [23] propose to distinguish between two spatial point clouds by measuring the Kullback–Leibler divergence between the probability density estimates of the samples. Similar density based techniques have been used in [22], [38], [39]. Formally, let $\mathcal{P}_P(x, y)$ and $\mathcal{P}_Q(x, y)$ be two probability density functions corresponding to the sets P and Q, which have been computed via parametric or non-parametric probability density estimation [40]. The Kullback–Leibler (KL) divergence based stability index of the set Q from P is defined as:

$$\rho_{KL} = \int \mathcal{P}_P(x, y) \log \left(\frac{\mathcal{P}_P(x, y)}{\mathcal{P}_Q(x, y)} \right) dxdy \qquad (2)$$

We would analyze the properties of this index in more details later, but we mention that for clustered detections, this index is capable of identifying the changes in keypoint distribution. However, for more homogeneous point distribution, the index in eq. (2) is less sensitive to local spatial deviation of the samples.

In a related topic, Bostanci *et al.* [34] propose a no-reference keypoint quality measure by analyzing the spatial coverage of the feature detections. In particular, the first order spatial statistic due to Ripley [20] is used to determine cluster properties of the point distributions. Feature coverage is an essential attribute for a keypoint detector [39], but this criteria cannot be considered as the sole metric for performance evaluation as it does not account for the stability to local variations or sensitivity to more global changes. Therefore, such no-reference metrics are typically used in conjunction with other reference based stability measures for an accurate representation of the efficacy of the keypoint detection system.

C. Motivation and outline of the proposed method

From the preceding discussions it becomes clear that the prevalent methods are either too local, or scale poorly to more global spatial variations. Another criticism of the correspondence seeking methodologies is that such techniques merely count the number of feature matches, and would report false associations (for randomly distributed target points) by disregarding their spatial distribution. Statistical analysis of such spatial point-sets is indeed essential for robust stability measurement. Stable detectors are largely unaffected by imaging artifacts and spatial sets (P and Q) of stable keypoints are expected to be strongly correlated. In contrast, spatial independence between the point-sets implies no statistically significant association, which is a strong indicator of instability. A close analogy is that of a binary classifier which randomly categorizes target data with probability 0.5. As the prior on the point-set distribution is unknown, complete spatial randomness (CSR) of the perturbed keypoints (realized via homogeneous Poisson process [21]) provides a suitable null hypothesis to test the degree and significance of association between the point-sets. Under this null hypothesis, association of perturbed keypoints to the reference data is statistically insignificant– an attribute of detector instability where imaging artifacts lead to unreliable and stochastically distributed keypoints.

Our proposed solution encapsulates these concepts in a two stage pipeline. First, a set of salient zones or cluster cores are computed based on the probability density function of the reference point cloud. This is defined in section III-B. Next we develop the mathematical framework to statistically estimate the degree of spatial correspondence of the perturbed keypoints to the cluster cores (see Fig. 1 for illustration) in section III-C, and subsequently define $\mathrm{C}^3\mathrm{I}$ as a measure of keypoint stability in section III-D. The technical details of our method are presented next.

III. Method

The first step of our algorithm is to define a set of salient zones on the image domain $\Omega \subset \mathbb{R}^2$, based on the spatial distribution of the reference set of keypoints $S_0 = \{\boldsymbol{\zeta}_i^0 \in \Omega\}, i \in [1, n_1]$. Intuitively, the cluster cores define a set of closed subregions which encapsulate the local clusters of the points in S_0, and they represent salient regions in the image with high local concentration of keypoints. Eventually, we are interested in finding the degree of correspondences of the perturbed keypoints in $S = \{\boldsymbol{\zeta}_j \in \Omega\}, j \in [1, n_2]$ to these cluster cores.

A simple strategy is to define the cluster cores as the union of local neighborhood regions $\cup_{i=1}^{n} b_r(\boldsymbol{\zeta}_i^0)$, where $b_r(\boldsymbol{\zeta}_i^0)$ is a r-neighborhood around each $\boldsymbol{\zeta}_i^0 \in S_0$. The drawback of this approach is that the parameter r is difficult to choose in practice. We present a method which is inspired by the dynamic clustering technique by Yip *et al.* [41]. A multi-scale methodology using deformable models is used to detect and represent the cluster core regions as the zero level set of an embedding function. This implicit representation using level sets [17] enables us to define the keypoint correspondence statistic for closed regions with arbitrary geometry. The mathematical details of our technique are presented in the following subsection.

A. Cluster core computation via level sets

Let us define a smooth, non-negative, and isotropic (2D) kernel function $K(\mathbf{z}) = e^{-\mathbf{z}^2}, \mathbf{z} \in \mathbb{R}^2$. For the reference set

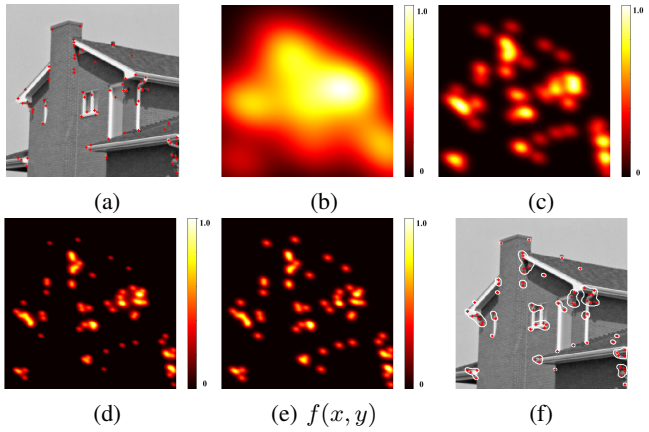


Fig. 2: FAST corner keypoints [45] on the *house* image are shown in red dots in (a). The kernel density estimates of the keypoints at different resolutions (s=1, 4, and 8) are shown in (b)-(d). The scale space density image via eq. 4 is shown in (e). The computed cluster core contours are displayed in white color in (f).

$S_0 = \{\boldsymbol{\zeta}_i^0\}$, the probability density function is computed for every pixel $\mathbf{x} \in \Omega$ as follows:

$$f_h(\mathbf{x}) = \frac{1}{n_1 h^2} \sum_{i=1}^{n_1} K\left(\frac{\mathbf{x} - \boldsymbol{\zeta}_i^0}{h}\right) \tag{3}$$

The kernel bandwidth h determines the smoothness of the estimated density function. Optimal choice of the bandwidth parameter is a challenging problem in pattern recognition. A popular method is due to Scott [40] who proposed a parameter selection strategy based on the effective number of points and the dimension of the data. Unfortunately, this has a tendency to estimate over-regularized solutions especially for multimodal distributions. Inspired by the scale-space medial axis detection technique in [42], we propose a multiscale approach to enhance the cluster core regions. If $\bar{h}$ is the bandwidth parameter estimated by Scott's technique, we define a scale-space as $\Theta = \{h_s\}$, where $h_s = \bar{h}/s$ and s is a positive integer. The corresponding density estimates at different scales are represented as $\{f_{h_s}\}_{s=1}^{p}$ where f_{h_s} is computed by substituting $h = h_s$ in eq. (3). Based on the multi-scale response, the optimal scale space kernel density estimate (KDE) response image is computed as follows:

$$f(\mathbf{x}) = \frac{1}{|\Theta|} \sum_{s \in \Theta} f_{h_s}(\mathbf{x}) \tag{4}$$

The multi-scale KDE image via eq. (4) is used to compute the cluster core contours using geometric active contours [43] by solving the following partial differential equation:

$$\frac{\partial \phi}{\partial t} = \text{div}\left(\left(\frac{1}{1 + |\nabla f|}\right) \frac{\nabla \phi}{|\nabla \phi|}\right) |\nabla \phi| \tag{5}$$

Here, $\phi : \Omega \mapsto \mathbb{R}$ is a Lipschitz function which assumes positive values inside its zero level contour, and negative outside. The initial condition of eq. (5) is defined by segmenting $f(\mathbf{x})$ using Otsu's technique [44]. The geodesic active contour framework defined in eq. (5) evolves the curves to produce smooth contours which are attracted to the high gradient regions of $f(\mathbf{x})$. Cluster cores are represented implicitly as the zero level set of the function ϕ. Fig. 2 illustrates the computation of the multiscale kernel density image. Fig. 2(b)-(d) show the density estimate obtained at coarser scales. The scale space response via eq. (4) is shown in (e), and the cluster core regions obtained via eq. (5) are shown in Fig. 2(f).

We acknowledge that besides the aforementioned solution, there exists other attractive techniques to extract such salient regions in an unsupervised fashion. An alternative would be to detect cluster cores using outlier detection methods such as isolation forests [46] and one-class SVM [47], or via density based clustering methods such as DBSCAN [48], although each method comes with their own set of intricacies. We prefer the proposed solution because it is a simple, yet elegant implicit representation of the cluster cores which is fully unsupervised. The segmentation model in eq. (5) is parameter free, and is automatically initialized by thresholding the multiscale density response using Otsu's technique [44]. Few iterations are required for convergence as the scale space KDE image already enhances the cluster cores and leads to precise contour initialization. The only free parameter in C^{3}I is the scale space in eq. (4), and a good initialization ($\bar{h}$) of the scale parameter is obtained using Scott's technique. The scale space is computed as $\Theta = \{\bar{h}, \ldots, h_s = \bar{h}/s\}$. Typically $s = 2^m$ ($m \in \mathbb{N}^+$), and we have observed experimentally that our algorithm is quite robust to the choice of this parameter for $3 \leq m \leq 5$. Sensitivity of C^{3}I to this parameter is reported in section IV-A, and Fig. 4(b) illustrates robustness of C^{3}I to parameter selection. Based on this experimental evidence we set $m = 4$ for all our experiments to optimize between speed and accuracy.

B. Spatial analysis of perturbed keypoints

The key ingredient of our stability measure is the ability to reject false keypoint correspondences due to random chance. As discussed earlier, instability is implied when the point-sets S and S_0 are spatially independent. This is mathematically realized by defining a null hypothesis where the spatial points in S are distributed according to the homogeneous Poisson law [21]. Homogeneous Poisson process is a special case of the generic Poisson point process (see Appendix for details) where the intensity function $\lambda : \Omega \mapsto \mathbb{R}^+$ is a positive scalar, *i.e.* $\lambda(x, y) = \lambda_0, \forall x, y \in \Omega$. Under this condition, the process mean simplifies to be $\mu(B) = \lambda_0 |B|$, where $B \subseteq \Omega$ is a closed sub-domain, and μ is a bounded Lebesgue measure. Mathematically, if N_B is a random variable which represents the number of events in B, the probability distribution of N_B is given as

$$\Pr\{N_B = k\} = \frac{e^{-\lambda_0 |B|} \left(\lambda_0 |B|\right)^k}{k!} \tag{6}$$

Homogeneous Poisson point processes are characterized by the special property of *complete spatial randomness* (CSR) i.e. the spatial location of a point is independent of another point in the process. Point samples from a homogeneous Poisson process are uniformly distributed over the sampling domain, which makes this an ideal candidate to model random spatial distribution of keypoints. With this model of randomness, we now define a correspondence statistic to the cluster core

regions, and analyze its characteristics when the keypoint distribution follows the homogeneous Poisson law.

C. Cluster core correspondence statistic

Let us denote the set of zero level contours of ϕ as $\mathcal{C} = \{\mathcal{C}_k\}$, where each non intersecting curve $\mathcal{C}_k$ encloses a region $\omega_k \subseteq \Omega$ and k is a non-negative integer. We further define a set membership indicator function $\mathbb{I}_S[\mathbf{y}] = 1$ if $\mathbf{y} \in S$ and zero otherwise.

Definition 1. *Given a set of closed hypersurfaces $\omega = \bigcup_k \omega_k$ and a set of spatial keypoints $S = \{\boldsymbol{\zeta}\}$, the cluster core correspondence statistic is defined as follows:*

$$\mathcal{K}_\omega = \int_{\mathbf{x}\in\Omega} \frac{\mathbb{I}_S[\mathbf{x}]}{\lambda(\mathbf{x})} H(\phi(\mathbf{x}))\, d\mathbf{x} \tag{7}$$

Here $\lambda(\mathbf{x})$ denotes the intensity function of the Poisson point process and $H(\phi)$ is the Heaviside function defined as $H(z) = 1$ if $z \geq 0$ and zero otherwise.

The cluster core correspondence statistic in eq. (7) is therefore proportional to the number of keypoint correspondences inside the cluster core regions obtained as $\omega = \{\mathbf{x} | \phi(\mathbf{x}) \geq 0\}$. We investigate eq. (7) when the set of perturbed keypoints $S = \{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_{n_2}\}$ follows a homogeneous Poisson process and analyze its statistical properties when the keypoint localizations are random. Under this assumption of complete spatial randomness, we substitute $\lambda(\mathbf{x}) = \lambda_0 = n_2/|\Omega|$ in eq. (7) and define a region indicator function χ_ω such that $\chi_\omega(\mathbf{y}) = 1$ if $\mathbf{y} \in \omega$ (or $H(\phi(\mathbf{y})) \geq 0$) and zero otherwise. Substituting these results in eq. (7), we simplify the statistic for the homogeneous case as follows:

$$\mathcal{K}_\omega = \frac{|\Omega|}{n_2} \sum_{j=1}^{n_2} \chi_\omega(\boldsymbol{\zeta}_j) \tag{8}$$

The statistic in eq. (8) is proportional to the number of points of the point process which are inside the cluster core region ω. The definition of $\mathcal{K}_\omega$ bears resemblance to the Ripley's K-function used in [34] which also counts the number of occurrences within a local neighborhood of the points. In fact, eq. (8) could be considered as a more generic form of the spatial statistic in [34] because it is defined for regions (ω) of arbitrary geometry. Additionally, the level set formulation does not require additional processing to analyze events near the domain boundary which makes the proposed statistic quite powerful and versatile.

Proposition 1. *If the set of spatial keypoints $S = \{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_{n_2}\}$ exhibits complete spatial randomness in Ω, $\mathcal{K}_\omega$ asymptotically converges to a Gaussian distribution.*

Proof. The proof follows from the analysis of the following random variable:

$$\chi = \sum_{j=1}^{n_2} \chi_\omega(\boldsymbol{\zeta}_j) \tag{9}$$

In (9), χ_ω is Bernoulli random variable, and therefore the sum of Bernoulli variables χ follows Binomial distribution (see (22) in Appendix for additional details) . Formally, $\chi \sim \texttt{Bin}(n_2, p_\omega)$ where $n_2 = |S|$, and the parameter of this distribution can be derived for homogeneous Poisson process as follows:

$$\begin{aligned} p_\omega = \mu(\omega)/\mu(\Omega) &= \frac{\int_\omega \lambda(x,y)dxdy}{\int_\Omega \lambda(x,y)dxdy} = \frac{1}{|\Omega|}\int_\Omega H(\phi)dxdy \\ &= |\omega|/|\Omega| \end{aligned} \tag{10}$$

We denote the area of the cluster core regions as $|\omega| = \int_\Omega H(\phi)dxdy$. To analyze the statistical characteristics of the correspondence statistic defined in (8) under CSR, we note that the asymptotic convergence of Binomial distribution towards the Normal distribution ensures that χ is a Gaussian random variable [49]. Since $\mathcal{K}_\omega$ is a scaled version of χ, it follows that $\mathcal{K}_\omega \sim \mathcal{N}(m_\omega, s_\omega)$, where m_ω and s_ω are the parameters of the distribution. The discussion for the convergence criteria closely follows the argument provided in [49] and further details are not furnished here. This property plays an important role in analysis since by virtue of the Gaussian distribution of the statistic, it is possible to characterize the clustering property of the point process if the parameters can be derived analytically.

Corollary 1. *Under the assumption of uniform random distribution of the spatial keypoints, the parameters of the normal distribution of the statistic $\mathcal{K}_\omega$ can be computed analytically to be $m_\omega = |\omega|$ and $s_\omega^2 = |\omega|(|\Omega| - |\omega|)/n_2$.*

Proof. It is straightforward to derive that the indicator function χ_ω is a Bernoulli random variable with expected value $\mathbb{E}[\chi_\omega] = p_\omega = |\omega|/|\Omega|$. Since the spatial points follow an uniform random distribution over Ω due to CSR, we proceed to analytically compute the mean and variance of the statistic $\mathcal{K}_\omega$ as follows:

$$\begin{aligned} m_\omega = \mathbb{E}[\mathcal{K}_\omega] &= \frac{|\Omega|}{n_2}\sum_{j=1}^{n_2}\mathbb{E}[\chi_\omega(\zeta_j)] \\ &= \frac{|\Omega|}{n_2}\sum_{j=1}^{n_2}|\omega|/|\Omega| = |\omega| \end{aligned} \tag{11}$$

Here $\mathbb{E}[.]$ represents the expected value of a random variable. Finally, we compute the variance as follows:

$$\begin{aligned} s_\omega^2 &= \mathbb{E}\left[(\mathcal{K}_\omega - m_\omega)^2\right] = \mathbb{E}\left[\mathcal{K}_\omega^2\right] - m_\omega^2 \\ &= \frac{|\Omega|^2}{n_2^2}\left(\sum_{m=1}^{n_2}\mathbb{E}\left[\chi_\omega^2(\boldsymbol{\zeta}_m)\right] + \right. \\ &\qquad \left. \sum_{m=1}^{n_2}\sum_{\substack{n=1 \\ n\neq m}}^{n_2}\mathbb{E}[\chi_\omega(\boldsymbol{\zeta}_m)\chi_\omega(\boldsymbol{\zeta}_n)] - n_2^2 p_w^2\right) \end{aligned} \tag{12}$$

Since χ_ω is a Bernoulli random variable, the term $\mathbb{E}\left[\chi_\omega^2(\boldsymbol{\zeta}_m)\right] = p_\omega$. Furthermore, the CSR property of the point process ensures that the spatial location of a point $\boldsymbol{\zeta}_m$ is independent of another point $\boldsymbol{\zeta}_n$, when $m \neq n$. Therefore, owing to this independence property, it can be further inferred

that $\mathbb{E}\left[\chi_\omega(\boldsymbol{\zeta}_m)\chi_\omega(\boldsymbol{\zeta}_n)\right] = \mathbb{E}\left[\chi_\omega(\boldsymbol{\zeta}_m)\right]\mathbb{E}\left[\chi_\omega(\boldsymbol{\zeta}_n)\right]$. Therefore, eq. (12) can be further simplified to

$$\begin{aligned} s_\omega^2 &= \frac{|\Omega|^2}{n_2^2}\left(n_2 p_\omega + n_2(n_2-1)p_w^2 - n_2^2 p_w^2\right) \\ &= |\omega|\left(|\Omega| - |\omega|\right)/n_2 \end{aligned} \tag{13}$$

From eq. (11) and eq. (13), we observe that the behavior spatial properties of the keypoint positions is fully characterized by the area of the cluster core regions $|\omega|$ and the number of points n_2. The area of the cluster cores is defined as $|\omega| = \int_\Omega H(\phi)dxdy$, and can be computed in the discrete domain using morphological connected component analysis. Therefore the parameters of the distribution of $\mathcal{K}_\omega$ can be calculated analytically without resorting to computationally intensive Monte Carlo techniques [50]. In the following section, we introduce the correspondence index which provides evidence of the amount of statistically significant coupling of the noisy keypoints to the reference.

D. Cluster Core Correspondence Index (C³I)

From Corollary 1, we obtain $\mathcal{K}_\omega \sim \mathcal{N}(m_\omega, s_\omega)$, and the parameters of the normal distribution are computed analytically in eq. (11) and eq. (13). Based on these analytic descriptions, we define a standardized statistic $\tilde{\mathcal{K}}_\omega \sim \mathcal{N}(0,1)$ as

$$\tilde{\mathcal{K}}_\omega = \left(\mathcal{K}_\omega - m_\omega\right)/s_\omega \tag{14}$$

A large positive value of $\tilde{\mathcal{K}}_\omega$ ($>> 0$) suggests evidence of accumulation of the keypoints to the cluster cores, whereas a smaller magnitude (≤ 0) indicates either a homogeneous spatial distribution, or significant dispersion of the keypoints. We further denote the variable $z \geq 0$ as

$$z = \tilde{\mathcal{K}}_\omega \mathbf{1}\left[\tilde{\mathcal{K}}_\omega\right] \tag{15}$$

where $\mathbf{1}[u] = 1$ if $u > 0$, and zero otherwise. The cluster core correspondence index $\rho_{C^3I} \in [0,1]$ between a reference S_0 and a point cloud set S is defined as:

$$\rho_{C^3I}\left[S_0||S\right] = \frac{1}{\beta}\kappa(z)s_\omega z \tag{16}$$

$$\text{where} \quad \kappa(z) = \sqrt{\frac{2}{\pi}}\int_{-\infty}^{z} e^{-\left(\frac{u^2}{2}\right)}du - 1 \tag{17}$$

The factor $\kappa \in [0,1]$ provides evidence of keypoint correspondence and $\kappa \approx 1$ when the keypoints are significantly coupled to the cluster cores, but assumes a lower value when the spatial distribution is more uniform. Here, β is a constant which normalizes the index to the range $[0,1]$. If we denote the pre-normalized index between a reference S_0 and a noisy set of measurements S to be $\rho\left[S_0||S\right] = \kappa s_\omega z$, the normalization factor is computed as $\beta = \rho\left[S_0||S_0\right]$.

E. Remarks

The fundamental aspect which distinguishes C³I from traditional evaluation studies is the manner in which keypoint stability is defined. Unlike most correspondence-based methods which essentially count the number of true feature matches within a pre-specified neighborhood, the mean m_ω and the variance s_ω of the correspondence statistic serve as a baseline to indicate clustering or dispersion of the detections. We argue that spatial randomness of the detected keypoints is a strong indicator of algorithm instability, although most existing evaluation techniques disregard this aspect. For example, if s_ω is finite, the variable z in eq. (16) is non-zero only when the number of correspondences in the cluster core region is statistically above the expected number of points under CSR. C³I scores favor the cases where $\mathcal{K}_\omega > m_\omega$, weighted by the confidence of accumulation κ. If the keypoint distribution is random, or if the keypoints are significantly dispersed from the cluster cores, $\mathcal{K}_\omega \leq m_\omega$, and consequently $\rho_{C^3I} = 0$. This distinguishes C³I from counting based measures such as eq. (1) which are not equipped to handle randomness. An example of such behavior of the different evaluation techniques is observed in Fig. 6. In the following section we present a set of experiments to compare and validate the proposed methodology against the state-of-the-art keypoint evaluation techniques.

IV. Experiments

In addition to the illustrations in Fig 2, in this section we provide experimental evidence to evaluate the robustness of keypoint detectors using both simulation studies and analysis on real-world datasets. This section is organized as follows. First, we discuss an experimental setup to quantitatively compare the performance of four different keypoint stability indices: (i) C³I (plotted in blue), (ii) the stability/repeatability index ρ_s (magenta) defined in eq. (1) [4], (iii) the KL-divergence index (green) [23], [37] ρ_{KL} defined in eq. (2), and finally (iv) the Mander's coefficient for binary images ρ_m (black) [18]. To compute ρ_{KL} we have used non-parametric kernel density method to estimate the spatial distribution, where the kernel bandwidth is selected using Scott's technique [40]. Since both ρ_m and ρ_s rely on the local radius parameter, we perform experiments with different values of r. This study (on both simulated test-cases and real-world examples) in illustrated in Sec. IV-A

In Sec. IV-B to IV-D, we further analyze how C³I is useful in assessing the robustness of keypoint detectors when the data is subject to varying degrees of imaging and measurement artifacts (such as noise, keypoint drift and in-plane rotation). In Sec. IV-E we use C³I to evaluate the performance of five state-of-the-art keypoint detection systems on an open-source natural image database under different imaging conditions. Further studies are presented in Sec. IV-F to show how C³I could also be interpreted as an effective measure of end-application (such as image matching, point cloud registration) performance evaluation. Finally, we demonstrate a few potential applications of C³I beyond keypoint stability analysis in Sec. IV-G. The following mathematical notations are used throughout this section. The reference set of keypoints (detected on non-corrupted data) is denoted by S_0, and the corresponding perturbed keypoint localizations are represented by the set S.

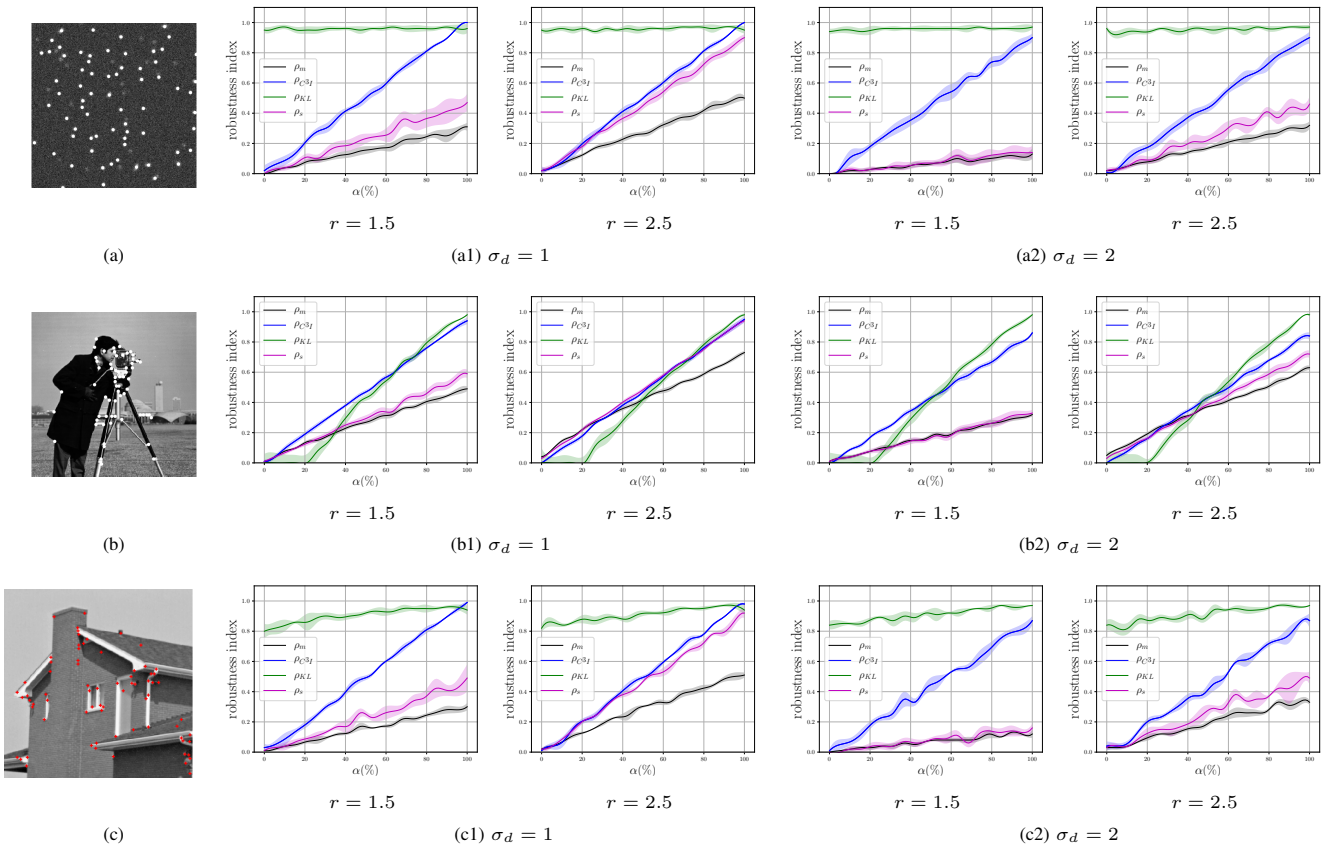


Fig. 3: Experimental analysis of correspondence estimation. (a)-(c) illustrates the keypoints extracted via Blob detector, ORB and FAST. The four robustness indices ρ_{C^3I}(blue), ρ_{KL}(green), ρ_s(magenta) and ρ_m(black) are plotted in (a1)-(c1) for $r = 1.5$ and $r = 2.5$ pixels and $\sigma_d = 1$ pixel. The corresponding results for $\sigma_d = 2$ pixels are presented in (a2)-(c2). The mean and standard deviation of the indices are shown, obtained via thirty Monte Carlo iterations for each value of the correspondence level α.

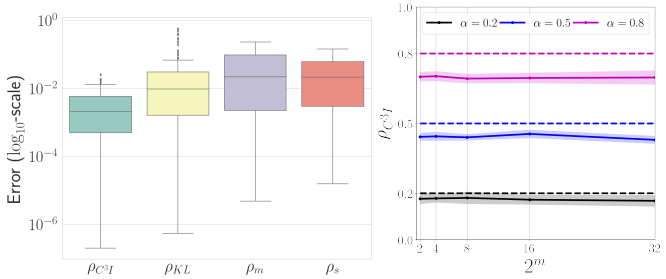
A. *Quantitative evaluation of stability indices*

Earlier we mentioned that statistical correspondence of keypoints provides insight regarding the stability of keypoint detectors, and robust descriptors are typically associated with high degree of association to the reference set. In this subsection we describe a set of experiments to study the behavior of the above mentioned correspondence indices. Three feature detectors are used to extract the keypoints - the wavelet based spot detector [28], Oriented and Rotated Brief (ORB) feature descriptor [13], and Harris' corner detector [51]. The reference detection due to the aforementioned detectors are shown in Fig. 3(a)-(c) respectively. We quantitatively analyze the efficiency of the above mentioned methods to estimate the degree of correspondence of the keypoints in S to the reference S_0. In this regard, we design a set of experiments where the elements in S are generated according to a Thomas process [21] to simulate spatial clustering of keypoints. Ideally, a stability index is expected to accurately infer the degree of correspondence of the perturbed set S to the reference, where S is a realization of a Thomas process with parameter $\alpha \in [0, 1]$ (coupling level, that indicates the level of association of the simulated spatial points to S_0). Therefore, an objective comparison of different stability indices may be performed by calculating the squared error between a stability index $\rho_{(.)}$ and the true correspondence level α as $e = |\rho_{(.)} - \alpha|^2$. For a statistically meaningful interpretation, average reconstruction error is reported over N independent Monte-Carlo trials.

The set S is generated via Thomas process simulation for a given $\alpha = \alpha'$ as follows. We first select a random set of $\alpha'|S_0|$ keypoints $\{\mathbf{p}\}$, and generate a set of perturbed locations $\{\mathbf{p}' = \mathbf{p} + \eta\}$ where $\eta \sim \mathcal{N}(0, \sigma_d)$ is a random variable. The remaining $(1 - \alpha')|S_0|$ keypoints are uniformly distributed over the image domain. The degree of association between the elements of S and S_0 correlates positively with the parameter α. Therefore, $\alpha = 0$ signifies spatially random keypoints, while $\alpha = 1$ simulates spatial keypoints which are significantly associated to S_0. Stability indices between S_0 and S are shown in Fig. 3(a1), (b1), and (c1) for $\sigma_d = 1$ pixels, and in Fig. 3(a2), (b2) and (c2) for $\sigma_d = 2$ pixels. Also, as mentioned earlier, the indices ρ_s and ρ_m rely on the radius parameter r, and we show the results for both $r = 1.5$ and $r = 2.5$ pixels as suggested in [3] and [18]. For our studies, α is sampled in $[0, 1]$ to obtain twenty discrete values, and we perform $N = 30$ iterations of Monte Carlo simulations for each α. The average robustness index and the associated variance are presented in Fig. 3. A qualitative analysis of the plots of the different correspondence indices in Fig. 3 suggests a largely linear trend for C^3I , whereas the other metrics are unable to capture the degree of association of to the desired level of precision. It can be observed that the indices ρ_s and

TABLE I: Mean squared error for correspondence estimation

		Mean Squared Error					
			$\rho_m(r)$		$\rho_s(r)$		
	Detector	ρ_{KL}	r = 1.5	r = 2.5	r = 1.5	r = 2.5	ρ_{C^3I}
$\sigma_d = 1$	Blob	0.29	0.16	0.07	0.09	0.003	**0.0001**
	ORB	0.009	0.081	0.02	0.058	**0.001**	**0.0008**
	FAST	0.22	0.16	0.04	0.10	0.0002	**0.0001**
$\sigma_d = 2$	Blob	0.29	0.25	0.15	0.24	0.09	**0.003**
	ORB	**0.008**	0.14	0.04	0.15	0.02	**0.008**
	FAST	0.23	0.26	0.13	0.24	0.08	**0.006**



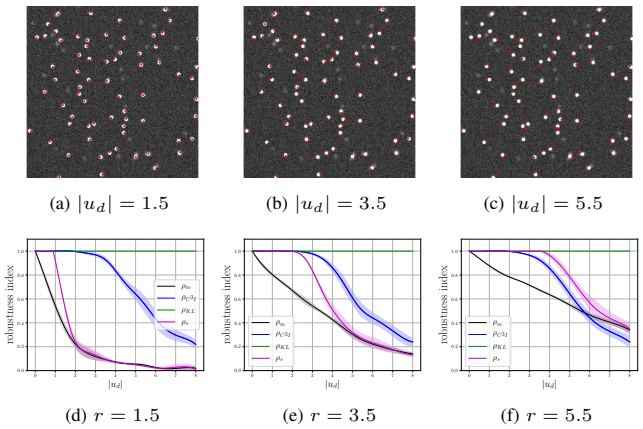
(a) (b)

Fig. 4: The boxplots in (a) show the distribution of (squared) errors for correspondence prediction on the de-raining dataset. The vertical axis is plotted in the $\log_{10}$ scale for better visualization. Parameter sensitivity analysis is shown in (b). Mean and standard deviation of C^{3}I response is plotted against the different values of the scale parameter m. the solid lines show C^{3}I correspondence estimated for three independent Thomas processes with $\alpha = 0.2$ (black), $\alpha = 0.5$ (blue) and $\alpha = 0.8$ (magenta) The dotted horizontal lines show the ideal correspondence values.

ρ_m are quite sensitive to the choice of the parameter r. The density based index ρ_{KL} exhibits a desirable linear trend when the feature detectors are clustered (see Fig. 3(b)), but it exhibits a significant presence of bias when the reference are more homogeneously distributed.

A quantitative description of the experiments is presented in Table-I, which reports the mean squared error ($MSE = \sum e/N$) of estimated correspondence computed over $N = 30$ Monte Carlo trials. The density based estimate ρ_{KL} is more applicable to situations where the reference detections are clustered, which is the case for ORB keypoints. An interesting aspect of this density based metric is that it is relatively immune to the local keypoint drift σ_d, although the performance lacks sensitivity when the spatial distribution is more homogeneous (for Blob detector and FAST). The quantitative evaluations also support the fact that both ρ_m and ρ_s rely heavily on the parameter r. Unfortunately, although in this case it may appear that increasing r could stabilize the results, we would demonstrate in the following experiments that the parameter needs to be chosen accurately for meaningful depiction of the point process. Finally, we find that C^{3}I outperforms the remaining techniques in terms of accuracy with the mean squared error less than 8×10^{-4} and 7×10^{-3} for $\sigma_d = 1$ and 2 respectively. This shows that C^{3}I is more adept at estimating the theoretical level of correspondence between interacting point processes, with significantly less error in correspondence estimation, except for the case of clustered ORB features where ρ_{KL} yields comparable performance.

We further perform additional experiments to analyze the performance of C^{3}I on a large open-source database of natural



(a) $|u_d| = 1.5$ (b) $|u_d| = 3.5$ (c) $|u_d| = 5.5$

(d) $r = 1.5$ (e) $r = 3.5$ (f) $r = 5.5$

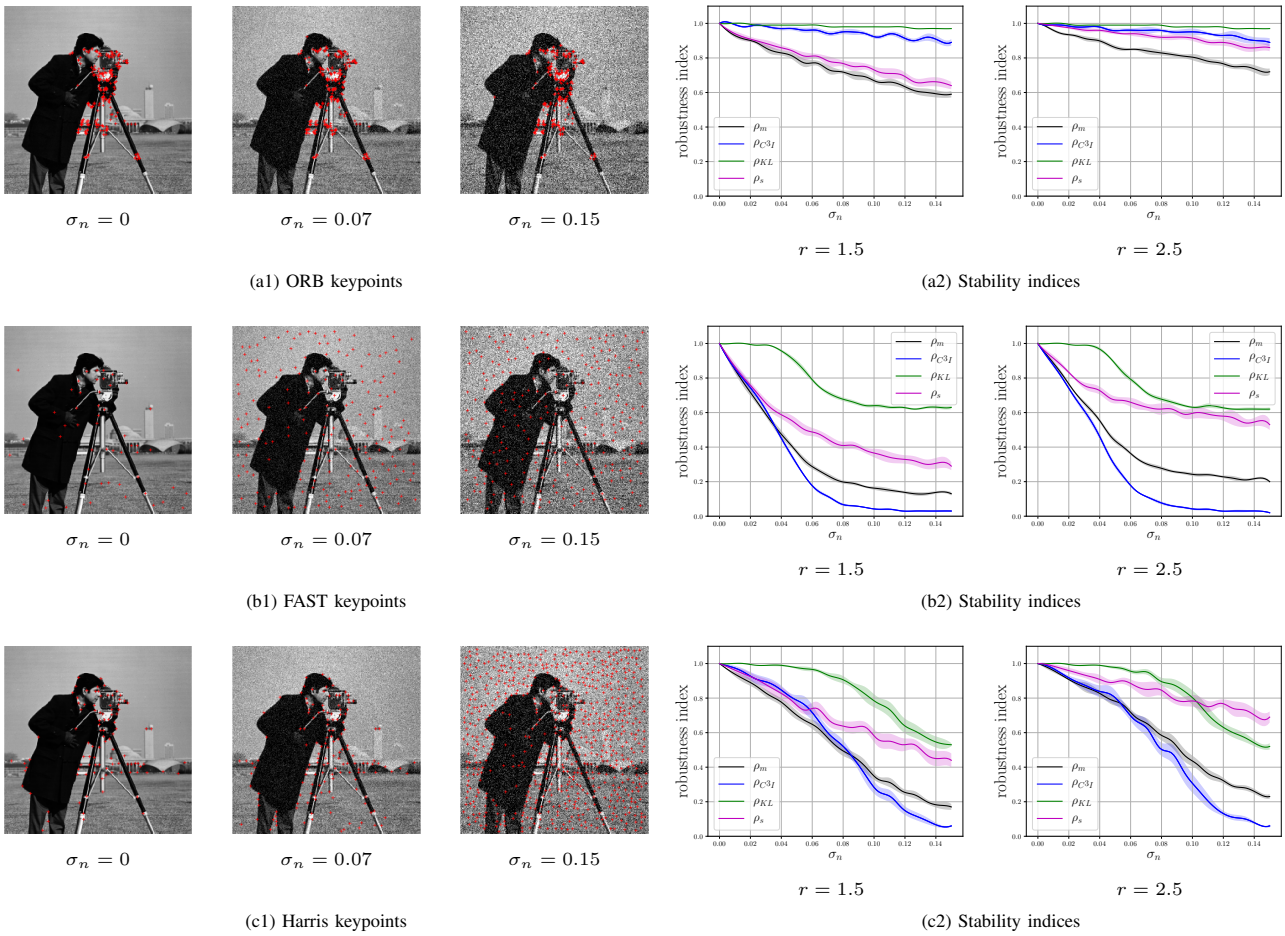
Fig. 5: Analysis of local feature drift. (a)-(c) shows the blobs detected via [28] in red due to a feature drift $u_d = \pm 1.5, \pm 3.5$ & $\pm$ 5.5 pixels. The reference set of keypoints are shown as white blobs. The four robustness indices ρ_{C^3I}(blue), ρ_{KL}(green), ρ_s(magenta) and ρ_m(black) are plotted in (d)-(f) corresponding to $r = 1.5, 3.5$ &5.5 pixels. The mean and standard deviation of the indices are shown, obtained via thirty Monte Carlo iterations for each value of u_d.

images [52]. One hundred images are selected at random, and the reference keypoints are computed on each image using the ORB algorithm for feature detection. The perturbed set is generated corresponding to each reference spatial dataset via Thomas process modeling as earlier, where the correspondence parameter α is chosen randomly between $[0, 1]$ for each image in the dataset. Thomas process noise parameter σ_d is set to 2 pixels and following the trend in the previous experiments, $r = 2.5$ is chosen for both ρ_s and ρ_m. Similar to the previous experimental setup, squared error in correspondence prediction is calculated for the four stability indices. The prediction error statistics are summarized by the boxplots (plotted in $\log_{10}$ scale) in Fig. 4(a), where the median error for each index is shown, and the box boundaries represent the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of the error distribution. The average squared error due to C^{3}I is calculated to be 4×10^{-3}, which is significantly better than the remaining three evaluation methods (ρ_{KL}, ρ_s, ρ_m) that result in average errors of 0.04, 0.05 and 0.03 respectively.

Finally, we perform experiments to justify the selection of the scaling factor $s = 2^m$ in eq. (4). We first select an image at random from the aforementioned dataset, and extract ORB keypoints for that image. The C^{3}I values are computed for $m = 2, 3, 4$ and 5, where the perturbed set is generated via Thomas process, with $\alpha = 0.2, 0.5$ and 0.8. The mean and variance of the C^{3}I response are shown in Fig. 4(b), computed over thirty Monte Carlo trials. This experimental data reveals that C^{3}I shows a stable trend in the range of experimental values of m, which makes it quite robust to its parameter.

B. Sensitivity to keypoint drift

The experiments in this section are performed to support our earlier claim that an ideal robustness index should be responsive to significant variation in keypoint position, while being robust to insignificant drifts. Although it is difficult to formally define what constitutes a significant change in localization, as a rule of thumb a spatial drift of less than three pixels is generally considered insignificant [4]. It is therefore



(a1) ORB keypoints (a2) Stability indices

(b1) FAST keypoints (b2) Stability indices

(c1) Harris keypoints (c2) Stability indices

Fig. 6: Qualitative and quantitative analysis of corner detectors to additive Gaussian noise. (a1)-(c1) illustrates the keypoints extracted on the cameraman image for three noise levels corresponding to ORB, FAST and Harris' detector. The four robustness indices ρ_{C^3I}(blue), ρ_{KL}(green), ρ_s(magenta) and ρ_m(black) are plotted in (a2)-(c2) corresponding to different noise levels, and for $r = 1.5$ and $r = 2.5$ pixels. The mean and standard deviation of the indices are shown, obtained via thirty Monte Carlo iterations for each value of σ_n.

expected of a robustness index to follow a smooth decaying trend with respect to the magnitude of drift in features when the change in position is no longer insignificant. For the experiments, we first compute a reference set of keypoints via the spot detector [28] on a simulated image of biological cells which was used for analysis in [49], [50]. Additionally, we define a drift factor u_d (pixels) and compute the set of drifted keypoints as $S = \{\mathbf{p} + \eta | \mathbf{p} \in S_0\}$, where $\eta \sim \mathcal{U}(-u_d, u_d)$. Here $\mathcal{U}(a, b)$ represents a uniform probability density function in $[a, b]$.

The results of the experiments are shown in Fig. 5. We perform experiments for $0 \leq |u_d| \leq 8.0$ pixels, and for each value of u_d, thirty Monte Carlo iterations are performed to compute the mean and variance of the stability indices. Fig. 5 (a)-(c) are provided to perform a visual inspection of the feature drifts for $u_d = 1.5, 3.5$ and 5.5 pixels. The reference keypoints S_0 are plotted as white blobs on the images, and the keypoints in S are plotted in red to show the difference in localization. It may be noted that the keypoint drifts are almost imperceptible in Fig. 5 (a), but it becomes more pronounced as u_d increases. The quantitative robustness indices are shown in Fig. 5 (d)-(f) corresponding to $r = 1.5, 3.5$ and 5.5 pixels for both ρ_s and ρ_m. We observe that while almost all the indices show a decaying trend, ρ_m and ρ_s are indeed affected by the choice of the radius. The index ρ_m tends to penalize even insignificant drifts, while ρ_s is more sensitive to the choice of the parameter r which effectively controls its sensitivity. Unfortunately, as previously discussed, selection of this parameter is non-trivial and this restricts the use of such indices for heterogeneous applications. Another observation is that the index ρ_{KL} does not show any perceptible change even when u_d is significant. This is because the detected spots are distributed homogeneously in the image, and therefore the spatial density function in eq. (3) does not show perceptible change even with local feature drifts. Interestingly, C^{3}I is less reliant on external parameters, and exhibits a smooth decaying trend which is in agreement with the visual perception of drift in feature localization.

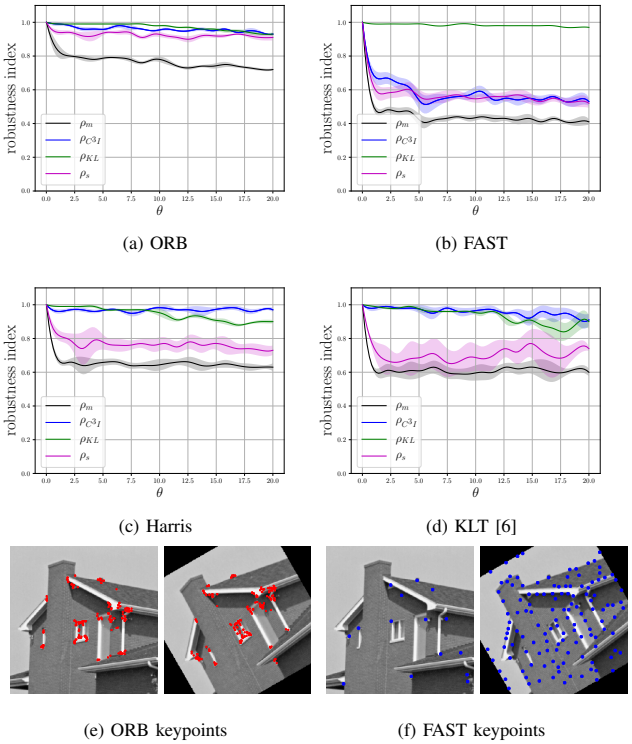
C. Sensitivity to noise

Noise is omnipresent in most practical image analysis applications, and noise invariance is a particularly desirable property of a keypoint detector. The objective of this set of experiments is to compute, and compare the performance of feature detectors in the presence of additive Gaussian noise. In particular, we evaluate three corner detectors: ORB [13],

FAST [45] and Harris' second-moment matrix based technique [51]. It is documented that Harris' method is very well suited to applications where noise is predominant, whereas FAST is known to be more sensitive to such artifacts. ORB is essentially an improved version of FAST which combines the accuracy of Harris' method with the speed of FAST. As earlier, the reference keypoints extracted from an image $h(x, y)$ is represented by the set S_0, and the set S comprises of the keypoints extracted from a noisy image $\tilde{h}(x, y) = h(x, y) + \eta$, where $\eta \sim \mathcal{N}(0, \sigma_n)$. We analyze the behavior of the above mentioned algorithms for $\sigma_n \in [0, 0.15]$ (image intensities are normalized to $[0, 1]$.). As earlier, we perform thirty Monte Carlo iterations for each value of σ_n and the mean and variance of the keypoint stability indices are shown in Fig. 6. As expected, the most stable performance is observed in ORB, which is well regarded for its robustness to additive noise. Remarkably, the algorithm is found to be stable to significant level of noise ($\sigma_n > 0.1$ pixels) and this could be verified both qualitatively in Fig. 6(a1), and quantitatively from Fig. 6(a2). A contrasting behavior is observed for FAST, which unlike ORB is found to be quite sensitive to noise both from a qualitative (Fig. 6(b1)) and quantitative perspective (Fig. 6(b2)). In fact, when the noise is significant, FAST keypoints tend to exhibit spatial randomness, which is reflected via $\rho_{C^3I} < 0.1$ for $\sigma_n > 0.07$. Here we make the observation that the response due to both ρ_{KL} and ρ_s appear to saturate even though the keypoint deviation is significant, although all four indices suggest an overall lack of stability of the algorithm to additive Gaussian noise. The results due to Harris' corner detection algorithm are shown in Fig. 6(c1)-(c2). While this technique is more robust compared to FAST, it is less adept than ORB at handling significant noise, resulting in a reduced stability index (less that 0.7) beyond a noise standard deviation of 0.07 pixels. This behavior can be explained by the fact that Harris-keypoints are detected at a single predefined scale which affects the stability for significant noise. This performance is verified qualitatively in Fig. 6(c1), where it is observed that a few noisy keypoints are extracted by the detector when the input is noisy, eventually yielding random localization when $\sigma_n > 0.14$ (see Fig. 6(c2)). We note that C^{3}I quantifies the expected stability of an algorithm which is verified qualitatively as well. The results suggest stability of ORB over Harris' technique and FAST, which is well known and has been reported in [11], [32]. An interesting phenomenon is observed when the spatial keypoints in S tend to be randomly spread for high noise levels as shown in Fig. 6(b1)-(b2) and Fig. 6(c1)-(c2). The plots of the different stability indices reveal that C^{3}I is capable of identifying system instability (especially when $\sigma_n > 0.1$), and penalizes such cases by reporting a significantly low value.

D. Assessment of rotational invariance

In this section, we describe a set of experiments to quantify the rotational invariance property of four popular corner detector algorithms, namely ORB, FAST [45], Harris' corner detector [51], and the KLT method [6]. Apart from FAST, the remaining three algorithms are known to be resilient to image rotations. FAST is a real time corner detection



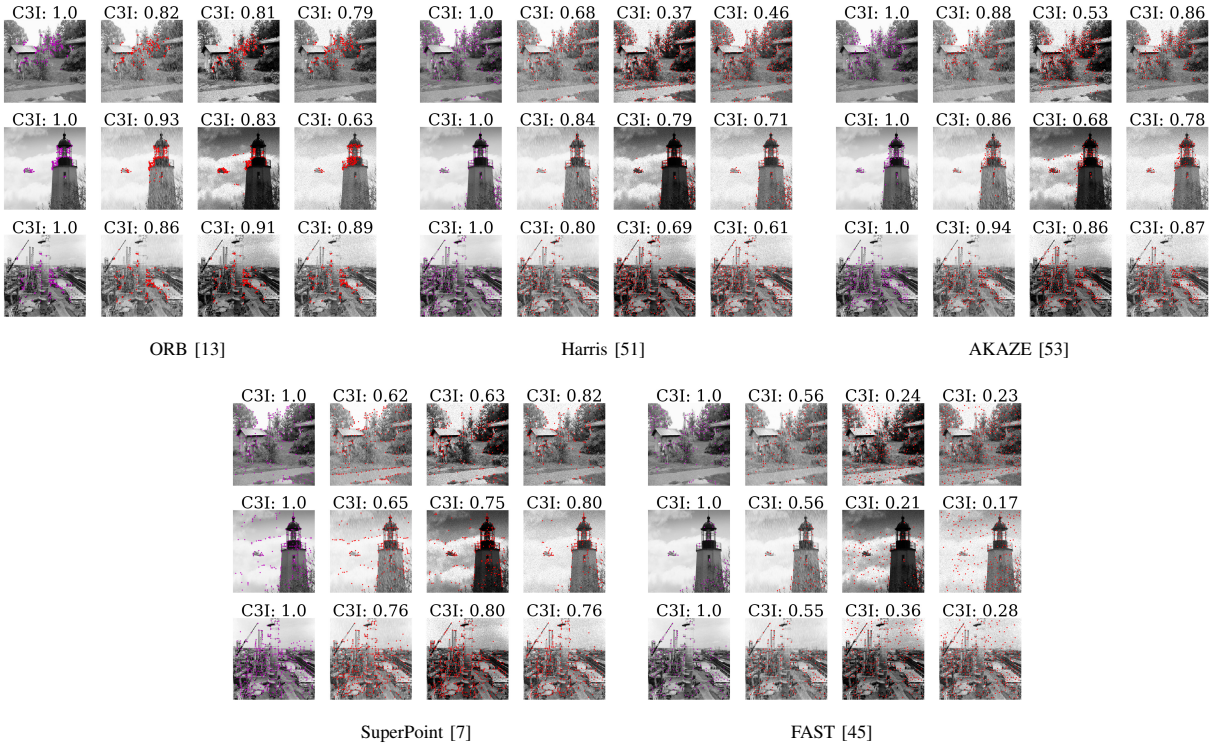
(a) ORB (b) FAST

(c) Harris (d) KLT [6]

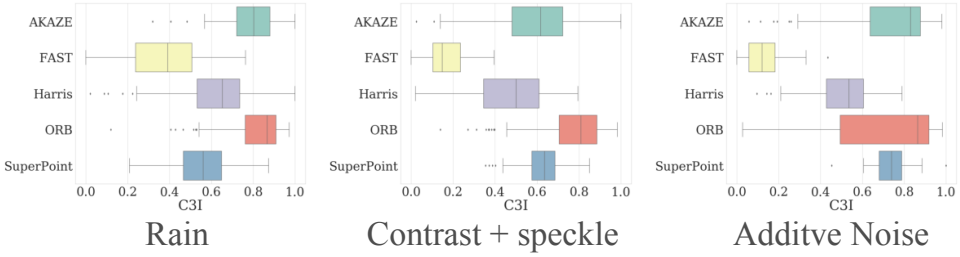
(e) ORB keypoints (f) FAST keypoints

Fig. 7: Analysis of rotational invariance of corner detectors (a) ORB, (b) FAST, (c) Harris, and (d) KLT. The four robustness indices ρ_{C^3I}(blue), ρ_{KL}(green), ρ_s(magenta) and ρ_m(black) are plotted for features extracted from an image rotated by $\theta \in [0, 20]$ degrees. The mean and standard deviation of the indices are shown, obtained via thirty Monte Carlo iterations for each value of θ. Sensitivity of FAST and ORB keypoints to in-plane rotation are illustrated in (e) and (f) respectively.

technique and it known to be sensitive to noise and geometric artifacts, a drawback which is later addressed [13]. These experiments are performed by extracting the reference set S_0 via a corner detector on the 'house' image shown in Fig. 3(c). Let us represent this image as $h(x, y)$ and denote $h_\theta(x, y)$ to be a rotated version of the image, θ signifying the degree of rotation. Let $\{\mathbf{p}_\theta\}$ denote the set of keypoints obtained on the rotated version of the image. We now compute the set $S = \{\mathcal{R}_\theta^{-1}\mathbf{p}_\theta\}$ where $\mathcal{R}_\theta$ is the rotation matrix. The quantitative analysis of the keypoint robustness indices are presented in Fig. 7. It is observed that ORB is most robust to rotations (see Fig. 7(a)), and the performance of Harris' technique and the KLT method are almost similar (Fig. 7(c and (d)). As expected, we observe a large change in the robustness index for FAST which is not particularly adept at handling geometric deformations (see Fig. 7(e)-(f) for comparison). In these studies we have set $r = 2.5$ for both ρ_m and ρ_s as prescribed in the literature. As earlier, for each value of θ, we perturb it by adding a Gaussian noise with standard deviation of five degrees, and perform thirty iterations of Monte Carlo simulations to compute the mean and variance of the robustness indices. It may be observed that while almost all four indices show a similar trend, ρ_s and more specifically ρ_m incorrectly suggest lack of robustness in Fig. 7(a), (c) and (d). This mostly due to the fact that the above mentioned indices are sensitive to insignificant change in keypoint localization, and therefore report a pessimistic quantification of the well



(a) Keypoint detection and corresponding C^3I evaluated on the de-raining dataset



(b) Boxplot of the C^3I scores for different artifacts

Fig. 8: Keypoint detections on three sample images from the dataset are shown in (a). Each detector is tested under the following degradation: rain artifacts, contrast change + speckle noise and additive Gaussian noise, and the detected (perturbed) keypoints are shown in red. The reference keypoints on the non-perturbed image is shown (in magenta) in the first columns. In (b), the performance of the keypoint detectors for the three different imaging artifacts (viz. rain, contrast, and additive noise) are summarized via boxplots.

known rotational invariant corner detectors.

E. Analysis of keypoint detectors on benchmark datasets

The following experiments are designed to show how C^3I could be effectively used to measure keypoint quality for natural images captured under varying external conditions. For this study, one hundred images are taken from the open source dataset provided by Zhang and Patel [52]. In addition to a variety of both indoor and outdoor reference scenes, this dataset also provides an additional set of images with appended *rain-streak* artifacts to simulate the effect of rain on digital captures. Such rain-streak effects degrade the image quality (SSIM = 0.71 ± 0.11) and present challenges for reliable keypoint estimation. We further augment the dataset by introducing contrast variations due to histogram equalization, and the image quality is further degraded by adding speckle noise to the data. This yields a dataset presenting significant illumination artifact with SSIM = 0.64 ± 0.11. Finally, a third set of perturbed data is generated by adding i.i.d Gaussian noise to the image pixels, yielding noise-corrupted image data with SSIM = 0.64 ± 0.07. Therefore, for every reference image in the dataset, three additional sets of artifact appended data is presented with varying degrees of image degradation. A montage of three representative images are shown in Fig. 8(a), where the first columns show the reference images, followed by corresponding effects introduced by rain artifacts, illumination variation and additive intensity noise.

We study the behavior of five state-of-the-art keypoint detectors under the aforementioned imaging conditions. In addition to the three feature detectors used before (Harris, FAST, and ORB), we further include two state-of-the-art techniques: AKAZE [53], and SuperPoint [7]. The latter is a semi-supervised deep convolutional neural network, pre-trained (implementation available on https://github.com/magicleap/SuperPointPretrainedNetwork.git) on a set of synthetically generated objects with varying levels of geometric and visual complexities. Keypoints identified by each detector are shown in Fig. 8(a), where the reference

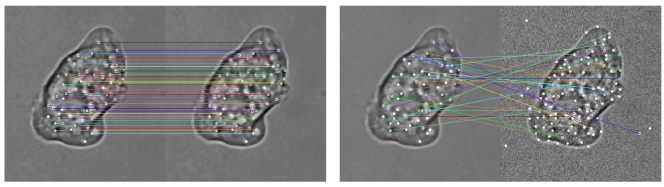
set of keypoints S_0 are plotted in magenta, and the red keypoints constitute the perturbed set S. Stability of the detection algorithm measured by C^{3}I are plotted for each image when subject to the different imaging conditions. A summary of the quantitative evaluation is presented in Fig. 8(b) via boxplots, where the box boundaries represent the 25^{th} and 75^{th} percentile C^{3}I scores, and the whiskers indicate the range of C^{3}I values. Statistically, ORB is found to be the most stable detector which yields the highest median C^{3}I score. The performance of ORB is closely followed by AKAZE, while FAST keypoints are found to be sensitive to significant changes in imaging conditions. Harris' corner points are also susceptible to lighting changes as the hessian based corner response function is known to be sensitive to intensity inhomogeneity.

It is specially interesting to analyze the semi-supervised SuperPoint algorithm. While statistically ORB and AKAZE are more stable than SuperPoint, it is noteworthy that the latter's stability could be significantly improved by fine tuning the neural network model on samples from this particular dataset. However, even when trained on a set of synthetic data, SuperPoint yields stable keypoints (with little variation in C^{3}I distribution). Indeed, the algorithm is less robust to rain-streak artifact, but this is expected because this particular implementation of SuperPoint is not supervised on data containing similar variability. However, it is quite interesting to note that traditional unsupervised detectors are still quite powerful from the perspective of keypoint stability and robustness.

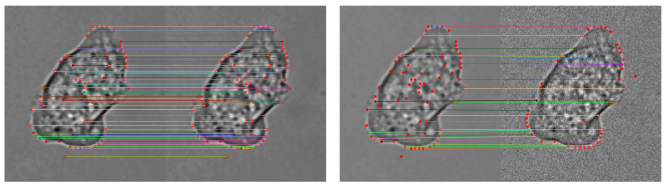
F. Correlating C^{3}I with downstream task performance

C^{3}I is a generic measure of keypoint quality, which enables validation of keypoint detectors irrespective of the deployed end-application. We hypothesize that robust keypoints are paramount to the success of any specific downstream task which uses the said detector in its pipeline. We present the following case studies to demonstrate this correlation between keypoint robustness (evaluated via C^{3}I) and the quality of the final solution.

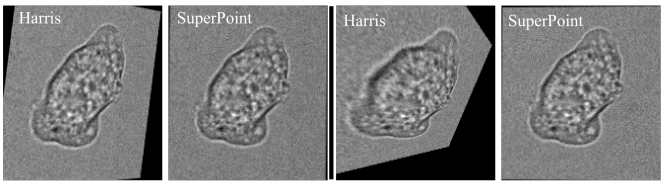
1) Keypoint based image alignment: Keypoint detectors are extensively used for finding correspondence between image pairs and to determine the homography matrix that defines the geometric transformation between the image-pairs. Such tasks require keypoints to be robust to imaging nuisances such as noise or contrast variations. The image pairs shown in Fig. 9 are geometrically related by an identity homography matrix, but reference image is a de-noised version of the noisy capture of an ameoba which is imaged via fluorescence microscopy. Keypoints are detected on both the reference image and its noisy counterpart due to Harris' corner detector in Fig. 9(a) and SuperPoint [7] detector in Fig. 9(b). The BRIEF [54] descriptor is associated with each spatial localization, and keypoint pairs are matched by a brute-force nearest-neighbor based matching technique. Correct correspondences (computed between same reference keypoints) are shown in the first column of Fig. 9(a) and (b). From the set of matched features, the RANSAC [55] algorithm is used to predict the affine transformation parameters. The alignment results of the noisy image to the reference are illustrated in Fig. 9(c). As



(a) Harris keypoint matching, reference-reference (left) vs reference-noisy (right)



(b) SuperPoint keypoint matching, reference-reference (left) vs reference-noisy (right)



(c) Estimated homography

Fig. 9: Effect of noise on feature matching and subsequent homography estimation. (a) Harris keypoints (shown as white dots) are matched both in absence and presence of noise, using brute force nearest-neighbor matching. Results for SuperPoint keypoints (red dots) are shown in (b), where the correspondences are more accurately recovered. In (c), feature based image alignment results are shown for both sets of keypoints using RANSAC for transformation estimation. Two such instances are shown to reflect the sensitivity of the final results on the keypoint quality.

RANSAC is stochastic in nature, two sets of predictions are shown. Since the actual homography matrix is identity, it becomes evident that SuperPoint-keypoints are more robust to noise than Harris' method. This end-result is validated from the C^{3}I scores which is significantly better for SuperPoint with $\rho_{C^3I} = 0.79$, compared to Harris' method ($\rho_{C^3I} = 0.33$) which explains the latter's sensitivity to intensity noise.

In addition to noise, contrast sensitivity is an important criteria while deciding on the appropriate detector for image alignment tasks. In Fig. 10, we further compare Harris' keypoints to the SuperPoint localizations for aligning image pairs with significant contrast variations. The HPatches database [56] contains images where the same scenery is imaged under different lighting conditions, as shown in Fig. 10 (top row). The first image (a) is captured under natural lighting conditions, and serves as the reference image for this experiment. Keypoints detected using Harris' corner response and SuperPoint method are shown in Fig. 10, with the reference keypoints S_0 plotted in the second row, and the corresponding keypoints S on the perturbed data are shown in the third row. The C^{3}I scores computed between the image pairs are presented as well. Significant contrast variation is known to affect traditional hessian-based corner detectors, and this is verified by a low stability score of $\rho_{C^3I}[a, c] = 0.56$. To find keypoint correspondence, the detected landmarks are first associated with the BRIEF descriptor, and then matches are found using a brute-force nearest neighbor search. The homography matrix relating the image pairs is then computed using RANSAC as earlier, and the final alignment results

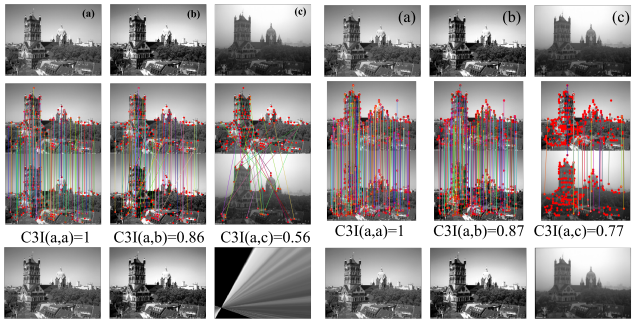


Fig. 10: Correlating C^{3}I with keypoint-based image alignment using RANSAC under contrast variations. Keypoint matches and the final alignment results are shown for Harris' keypoints and SuperPoint.

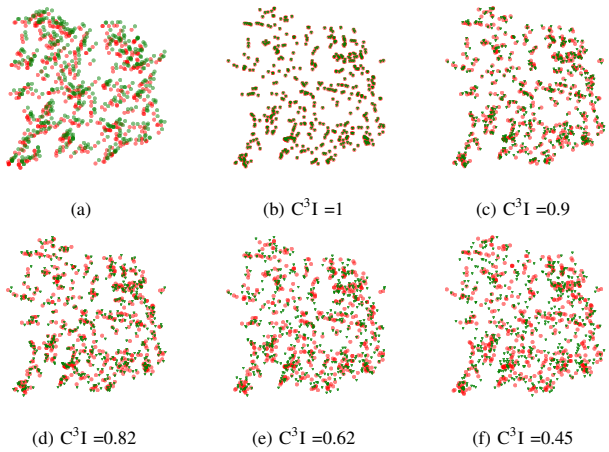


Fig. 11: Correlating C^{3}I values with CPD-based point cloud matching performance for different levels of local perturbations u_d of the spatial localizations. Initial spatial locations of the two different point clouds are shown in (a). The reference point cloud is plotted in red. The green points are perturbed locally by u_d pixels, and the final registration results (green markers) are shown in (b)-(f) corresponding to progressively higher values of $u_d = \{0, 1.5, 2.0, 3.0, 4.0\}$ pixels respectively. The C^{3}I scores are found to be in agreement with the efficacy in registering the point-clouds.

are shown in the bottom row of Fig. 10(a) and (b). The error in alignment is most prominent in the third example in Fig. 10(a) which is caused by the contrast sensitivity of the Harris-keypoints. The final image alignment results in Fig. 10(bottom row) strongly agree with the C^{3}I scores, and the RANSAC based alignment estimation is only slightly affected by the SuperPoint method which is more stable to such illumination artifacts. This provides further evidence to connect the keypoint stability measure C^{3}I with the end-product efficacy for image alignment operations.

2) Point-cloud registration using Coherent Point Drift: The keypoint stability measure proposed in this work also extends to unstructured spatial objects such as point clouds. Here, we consider a point cloud of spatial coordinates of a *Drosophila* neuron obtained from an open-source digital atlas of neuron tracings: *NeuroMorpho* [57]. The spatial points are sub-sampled, and projected on a 100×100 discrete grid, shown in red color in Fig. 11(a). We consider a second set of points which is translated from the red points by 2 units, and this translated set is plotted in green in Fig. 11(a). Our objective is to register the green points to the red samples using the Coherent Point Drift (CPD) algorithm [26]. Indeed, when the green points are unperturbed ($\rho_{C^3I} = 1$) the points are related by simple translation, and can be registered using CPD with little error (see Fig. 11(b)). However, registration is non-trivial when the spatial coordinates of the green points $\mathbf{x}_g$ are perturbed as $\mathbf{x}'_g = \mathbf{x}_g + \eta$, where $\eta \sim \mathcal{U}(-u_d, u_d)$ units. Such random perturbation often occurs in practice due to uncertainty in localizing the individual points. Effects of this uncertainty in keypoint localization are reflected in the registration performance, which gradually worsens for higher values of u_d. This phenomena is well captured by the gradual lowering of C^{3}I , illustrated in Fig. 11(b)-(f). This further supports our hypothesis that keypoint quality measure via C^{3}I is a good indicator of downstream performance.

G. Other potential applications and perspectives

Here we present two potential use cases outside the domain of keypoint stability quantification where C^{3}I could be successfully integrated. The first application involves using C^{3}I as a loss function for registering point clouds. We demonstrate a few experimental cases to illustrate how C^{3}I could be useful as a registration metric to be used in conjunction with other similarity measures such as the minimum mean squared error or KL divergence between the point cloud densities. Another application is described, where we extend C^{3}I to understand the similarity between different feature detectors.

1) C^{3}I as a loss function for point-cloud registration: We consider two special cases where we compare two different data distributions. The first example illustrated in the first row of Fig. 12, shows two point clouds $X \sim \mathcal{N}(0, \Sigma)$ (plotted in green) and a rotated set $X_\theta \sim \mathcal{N}(0, \Sigma_\theta)$ (in blue). Here θ is the rotation parameter, Σ_θ is the covariance matrix for the rotated dataset which is computed as follows:

$$\Sigma_\theta = \mathcal{R}_\theta^t \Sigma \mathcal{R}_\theta \tag{18}$$

Here $\mathcal{R}_\theta$ is the standard rotation matrix. We consider X to be the reference dataset, and the cluster cores are computed based on this distribution (shown as the contour line in black). We perform experiments by sampling 500 points from each distribution for $\theta \in [0, \pi]$ and analyze the effect of rotation of data on the C^{3}I . For sensitivity analysis, observations are sampled thirty times for each value of θ, and the average C^{3}I index is plotted in the last column (top row) along with the standard deviation. The plot of ρ_{C^3I} shows a monotonic trend, where the response is maximum when the two datasets are optimally aligned (for $\theta = 0$ or $\theta = \pi$). Similarly, we also analyze the response of C^{3}I to translation, where the translated data are assumed to be a realization of $X_d \sim \mathcal{N}(\mu_d, \Sigma)$ where $\mu_d = [0, d]$ is the translated mean and $d \in [-10, 10]$ units. The plot of ρ_{C^3I} increases monotonically as the point-cloud separation approaches zero, which corresponds to the peak response. Next, we perform similar experiments on a real-world 2D point cloud data of Drosophila neuron reconstructions available from the online repository: *www.neuromorpho.org*. The target point set $\mathcal{P} = \{\mathbf{x}_p\}$ is created by projecting the discrete neuron traces to a 100×100 grid (shown in red color in Fig. 12(c)). We further compute a second set of source points (blue) $\mathcal{Q} = \{\mathbf{x}_q\}$ related to the target set by the following

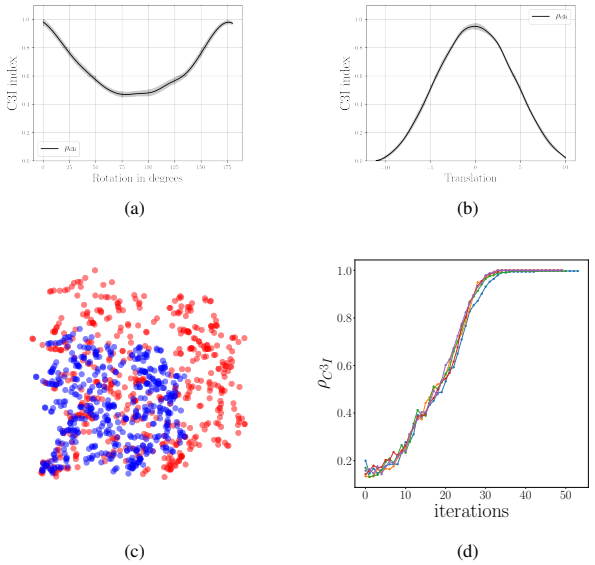


Fig. 12: C^3I scores are shown for registering rotated and translated 2D Gaussian data in (a) and (b) respectively. The target set (red) and the source set (blue) for point-set registration are illustrated in (c). In (d), C^3I values during the successive iterations of CPD algorithm is plotted for five distinct registration experiments.

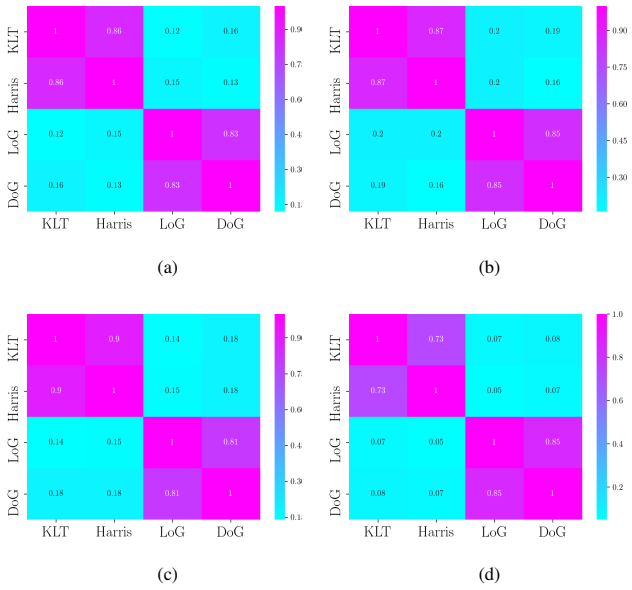


Fig. 13: The confusion matrices illustrate the keypoint similarity via the symmetric C^3I metric. Two corner detectors (KLT and Harris) are compared against two blob detectors (LoG and DoG) for four imaging scenarios corresponding to noise level $\sigma_n = 0, 0.01, 0, 03$ and 0.1 in (a)-(d) respectively. The quantitative metric reveals high intra-class and low inter-class similarity.

affine relation $\mathbf{x}_q = \mu\mathbf{x}_p + \gamma$, where $\mu \sim \mathcal{U}[0.6, 1.2]$ and $\gamma \sim \mathcal{U}[0, 3.0]$. A representative scatter-plot of the point sets is presented in Fig. 12(c). Our objective is to register the source points to the target set using the CPD [26] algorithm for point set registration. CPD is an iterative technique to find both rigid and non-rigid deformations between point clouds using a density estimation strategy. At each iteration of CPD, we record the C^3I values between the target points $\mathcal{P}$ and the deformed source point-set. The experiment is performed five times and the C^3I responses are plotted for each CPD iteration in Fig. 12(d). It is observed that over the iterative procedure, the CPD algorithm maximizes the C^3I response between the two point clouds while seeking the optimal deformation parameters. While these experiments are preliminary, the results indicate that C^3I could be used to identify the statistical deviation between point clouds for point cloud registration problems. Following the recent trend in using deep learning to estimate the registration parameters, it would be a future work to integrate C^3I as a differentiable loss function in learning based algorithms [25], [58] which could be used either by itself, or in conjunction with other measures of point cloud similarity.

2) C^3I as a keypoint similarity measure: The final set of studies are designed to illustrate an application of C^3I which would be useful to understand the similarity (or the difference) between different feature detectors. This could be useful to characterize and cluster algorithms which exhibit similar characteristics. We show via examples that C^3I could be used to understand the relationship between two or more feature detectors. We use four keypoint detector algorithms for this proof-of-concept study, which consists of two corner detectors (Harris and KLT) and two blob detectors (*Laplacian of Gaussian* or LoG and *Difference of Gaussian* or DoG) filters [1]. We first define a symmetric similarity index between the keypoint sets S_1 and S_2 to be

$$\rho^{'}_{C^3I} = \frac{1}{2}\left(\rho_{C^3I}[S_1||S_2] + \rho_{C^3I}[S_2||S_1]\right) \tag{19}$$

In Fig. 13, the symmetric C^3I index in eq. (19) is shown when the keypoints extracted by via one detector (say LoG) is compared to the keypoint localization due to a second detector (say ORB). The responses are plotted in Fig. 13(a)-(d) for different levels of additive noise, and we observe an average within-class similarity measure > 0.82 while the average inter class similarity is less than 0.15. This strongly indicates that C^3I is able to quantify the similarity between algorithms, which could be used to cluster the feature detection techniques into different semantic categories (corners and blobs for this specific case). This is more crucial in understanding the behavior of some end-to-end learning based detectors by associating the more abstract, yet rich keypoints to the low level cues in the image, such as blobs, corners, junctions, etc. This would be beneficial to further our knowledge about the intrinsic behavior of such algorithms which are unfortunately often used as a black box, thereby allowing users to choose the more appropriate technology for a specified task.

V. Conclusion

In this paper we have introduced a technique which builds on the theory of spatial pattern analysis and quantifies a detector's robustness by computing statistical correspondence of keypoints to the cluster core regions of the reference. The efficacy of C^3I has been established via both quantitative and qualitative assessment on various experimental studies which suggest a wide applicability of the solution. C^3I is application agnostic, and could be used universally as a test-bench to evaluate the robustness and stability of feature detectors. A

future extension to the presented study would be to detect feature correspondence in a high dimensional space, such as the intermediate projection subspace for deep autoencoders. While C^{3}I is designed specifically for spatial keypoints (in $\mathbb{R}^2$), extension to a higher dimension will be possible provided a suitable distance metric for such instances. This is indeed an exciting research prospect which will be discussed in more detail in our subsequent research.

APPENDIX

The mathematical preliminaries of a Poisson point process is presented here. Let $\lambda : \Omega \mapsto \mathbb{R}$ denote an integrable intensity function on the domain Ω which represents the number of events per unit area. Based on this definition, a bounded Lebesgue measure μ can be defined on any subset $A \subseteq \Omega$ as follows:

$$\mu(A) = \int_A \lambda(\mathbf{x}) d\mathbf{x} \tag{20}$$

Now let us consider a singleton point process $\{\boldsymbol{\zeta}\}$ which is characterized by a probability density function $f(\boldsymbol{\zeta}) = \lambda(\boldsymbol{\zeta})/\mu(\Omega)$. It is straightforward to verify that $\int_\Omega f(\boldsymbol{\zeta}) d\boldsymbol{\zeta} = 1$. Therefore, for the region A, we can define the probability p_A for $\boldsymbol{\zeta} \in A$ as follows:

$$p_A = \Pr\{\boldsymbol{\zeta} \in A\} = \int_A f(\boldsymbol{\zeta}) d\boldsymbol{\zeta} = \frac{\mu(A)}{\mu(\Omega)} \tag{21}$$

If N_A denotes the number of events in A, the conditional probability of N_A(conditioned on $N_\Omega = |S|$) can be expressed via the binomial law as follows:

$$\Pr\{N_A = k | N_\Omega = |S|\} = \binom{|S|}{k} p_A^k \left(1 - p_A\right)^{|S|-k} \tag{22}$$

According to eq. (22), the *i.i.d.* spatial point set S defines a binomial process over the space Ω conditioned on the cardinality of the point set. When $|S|$ is not specified, the spatial points follow a Poisson point process [20] described as $N_B \sim \texttt{Poisson}\left(\mu\left(B\right)\right)$, where $\mu(B) < \infty$ and B is any bounded subset of Ω. Formally,

$$\Pr\{N_B = k\} = \frac{e^{-\mu(B)} \mu(B)^k}{k!} \tag{23}$$

Here the mean of the Poisson process $\mu(B)$ denotes the expected number of points in B.

REFERENCES

[1] K. Grauman and B. Leibe, "Visual object recognition," *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 5, no. 2, pp. 1–181, 2011.

[2] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[3] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[4] ——, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.

[5] S. Gauglitz, T. Höllerer, and M. Turk, "Evaluation of interest point detectors and feature descriptors for visual tracking," *International Journal of Computer Vision*, vol. 94, no. 3, p. 335, 2011.

[6] J. Shi and C. Tomasi, "Good features to track," Cornell University, Tech. Rep., 1993.

[7] D. DeTone, T. Malisiewicz, and A. Rabinovich, "Superpoint: Self-supervised interest point detection and description," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 224–236.

[8] R. Sarkar and S. T. Acton, "Sdl: Saliency-based dictionary learning framework for image similarity," *IEEE Transactions on Image Processing*, vol. 27, no. 2, pp. 749–763, 2017.

[9] J. Li and G. H. Lee, "Usip: Unsupervised stable interest point detection from 3d point clouds," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 361–370.

[10] G. Tinchev, A. Penate-Sanchez, and M. Fallon, "Skd: Unsupervised keypoint detecting for point clouds using embedded saliency estimation," *arXiv preprint arXiv:1912.04943*, 2019.

[11] J. Hartmann, J. H. Klüssendorff, and E. Maehle, "A comparison of feature descriptors for visual slam," in *2013 European Conference on Mobile Robots*. IEEE, 2013, pp. 56–61.

[12] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[13] E. Rublee, V. Rabaud, K. Konolige, and G. R. Bradski, "Orb: An efficient alternative to sift or surf." in *ICCV*, vol. 11, no. 1, 2011, p. 2.

[14] A. Gordo, J. Almazan, J. Revaud, and D. Larlus, "End-to-end learning of deep visual representations for image retrieval," *International Journal of Computer Vision*, vol. 124, no. 2, pp. 237–254, 2017.

[15] N. A. Thacker, A. F. Clark, J. L. Barron, J. R. Beveridge, P. Courtney, W. R. Crum, V. Ramesh, and C. Clark, "Performance characterization in computer vision: A guide to best practices," *Computer Vision and Image Understanding*, vol. 109, no. 3, pp. 305–334, 2008.

[16] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli *et al.*, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[17] S. Mukherjee and S. T. Acton, "Region based segmentation in presence of intensity inhomogeneity using legendre polynomials," *IEEE Signal Processing Letters*, vol. 22, no. 3, pp. 298–302, 2014.

[18] J. Chen, L.-h. Zou, J. Zhang, and L.-h. Dou, "The comparison and application of corner detection algorithms." *Journal of Multimedia*, vol. 4, no. 6, 2009.

[19] F. Tombari, S. Salti, and L. Di Stefano, "Performance evaluation of 3d keypoint detectors," *International Journal of Computer Vision*, vol. 102, no. 1-3, pp. 198–220, 2013.

[20] P. J. Diggle *et al.*, *Statistical analysis of spatial point patterns*. Academic press, 1983.

[21] T. Lagache, A. Grassart, S. Dallongeville, O. Faklaris, N. Sauvonnet, A. Dufour, L. Danglot, and J.-C. Olivo-Marin, "Mapping molecular assemblies with fluorescence microscopy and object-based spatial statistics," *Nature communications*, vol. 9, no. 1, p. 698, 2018.

[22] M. Xie, J. Hu, S. Guo, and A. Y. Zomaya, "Distributed segment-based anomaly detection with kullback–leibler divergence in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 101–110, 2016.

[23] F. Järemo Lawin, M. Danelljan, F. Shahbaz Khan, P.-E. Forssén, and M. Felsberg, "Density adaptive point set registration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3829–3837.

[24] G. C. Sharp, S. W. Lee, and D. K. Wehe, "Icp registration using invariant features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 90–102, 2002.

[25] K. Zampogiannis, C. Fermuller, and Y. Aloimonos, "Topology-aware non-rigid point cloud registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[26] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[27] Z. Püspöki and M. Unser, "Template-free wavelet-based detection of local symmetries," *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3009–3018, 2015.

[28] J.-C. Olivo-Marin, "Extraction of spots in biological images using multiscale products," *Pattern Recognition*, vol. 35, no. 9, pp. 1989–1996, 2002.

[29] E. Simo-Serra, E. Trulls, L. Ferraz, I. Kokkinos, P. Fua, and F. Moreno-Noguer, "Discriminative learning of deep convolutional feature point descriptors," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 118–126.

[30] A. B. Laguna, E. Riba, D. Ponsa, and K. Mikolajczyk, "Key. net: Keypoint detection by handcrafted and learned cnn filters," *arXiv preprint arXiv:1904.00889*, 2019.

[31] Y. Verdie, K. Yi, P. Fua, and V. Lepetit, "Tilde: a temporally invariant learned detector," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 5279–5288.

[32] F. Mokhtarian and F. Mohanna, "Performance evaluation of corner detectors using consistency and accuracy measures," *Computer Vision and Image Understanding*, vol. 102, no. 1, pp. 81–94, 2006.

[33] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *International Journal of Computer Vision*, vol. 37, no. 2, pp. 151–172, 2000.

[34] E. Bostanci, N. Kanwal, and A. F. Clark, "Spatial statistics of image features for performance comparison," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 153–162, 2014.

[35] K. W. Dunn, M. M. Kamocka, and J. H. McDonald, "A practical guide to evaluating colocalization in biological microscopy," *American Journal of Physiology-Cell Physiology*, vol. 300, no. 4, pp. C723–C742, 2011.

[36] T. Hong-Phuoc and L. Guan, "A novel key-point detector based on sparse coding," *IEEE Transactions on Image Processing*, vol. 29, pp. 747–756, 2019.

[37] W. Zhang, L. Wenyin, and K. Zhang, "Symbol recognition with density matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2020–2024, 2006.

[38] A. Giantomassi, F. Ferracuti, S. Iarlori, G. Ippoliti, and S. Longhi, "Electric motor fault detection and diagnosis by kernel density estimation and kullback–leibler divergence based on stator current measurements," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1770–1780, 2014.

[39] G. Dán, M. A. Khan, and V. Fodor, "Characterization of surf and brisk interest point distribution for distributed feature extraction in visual sensor networks," *IEEE Transactions on Multimedia*, vol. 17, no. 5, pp. 591–602, 2015.

[40] D. W. Scott, *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons, 2015.

[41] A. M. Yip, C. Ding, and T. F. Chan, "Dynamic cluster formation using level set methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, pp. 877–889, 2006.

[42] S. Mukherjee and S. T. Acton, "Vector field convolution medialness applied to neuron tracing," in *International Conference on Image Processing*. IEEE, 2013, pp. 665–669.

[43] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[44] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[45] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *European Conference on Computer Vision*. Springer, 2006, pp. 430–443.

[46] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *IEEE International Conference on Data Mining*. IEEE, 2008, pp. 413–422.

[47] B. Schölkopf, R. C. Williamson, A. J. Smola, J. Shawe-Taylor, and J. C. Platt, "Support vector method for novelty detection," in *Advances in Neural Information Processing Systems*, 2000, pp. 582–588.

[48] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "Dbscan revisited, revisited: why and how you should (still) use dbscan," *ACM Transactions on Database Systems (TODS)*, vol. 42, no. 3, pp. 1–21, 2017.

[49] T. Lagache, V. Meas-Yedid, and J.-C. Olivo-Marin, "A statistical analysis of spatial colocalization using ripley's k function," in *International Symposium on Biomedical Imaging*. IEEE, 2013, pp. 896–901.

[50] T. Lagache, N. Sauvonnet, L. Danglot, and J.-C. Olivo-Marin, "Statistical analysis of molecule colocalization in bioimaging," *Cytometry Part A*, vol. 87, no. 6, pp. 568–579, 2015.

[51] C. G. Harris, M. Stephens *et al.*, "A combined corner and edge detector." in *Alvey vision conference*, vol. 15, no. 50. Citeseer, 1988, pp. 10–5244.

[52] H. Zhang and V. M. Patel, "Density-aware single image de-raining using a multi-stream dense network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 695–704.

[53] P. F. Alcantarilla and T. Solutions, "Fast explicit diffusion for accelerated features in nonlinear scale spaces," *IEEE Trans. Patt. Anal. Mach. Intell*, vol. 34, no. 7, pp. 1281–1298, 2011.

[54] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: Binary robust independent elementary features," in *European Conference on Computer Vision*. Springer, 2010, pp. 778–792.

[55] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[56] V. Balntas, K. Lenc, A. Vedaldi, and K. Mikolajczyk, "Hpatches: A benchmark and evaluation of handcrafted and learned local descriptors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5173–5182.

[57] G. A. Ascoli, D. E. Donohue, and M. Halavi, "Neuromorpho. org: a central resource for neuronal morphologies," *Journal of Neuroscience*, vol. 27, no. 35, pp. 9247–9251, 2007.

[58] Y. Wang and J. M. Solomon, "Deep closest point: Learning representations for point cloud registration," *arXiv preprint arXiv:1905.03304*, 2019.