



Harris Corners in the Real World: A Principled Selection Criterion for Interest Points Based on Ecological Statistics

Neil Bruce, Pierre Kornprobst

► To cite this version:

Neil Bruce, Pierre Kornprobst. Harris Corners in the Real World: A Principled Selection Criterion for Interest Points Based on Ecological Statistics. [Research Report] RR-6745, INRIA. 2008, pp.22. inria-00343449

HAL Id: inria-00343449

<https://hal.inria.fr/inria-00343449>

Submitted on 1 Dec 2008

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.



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

*Harris Corners in the Real World: A Principled
Selection Criterion for Interest Points Based on
Ecological Statistics*

Neil Bruce — Pierre Kornprobst

N° 6745

December 2008

Thème BIO

*R*apport
de recherche



Harris Corners in the Real World: A Principled Selection Criterion for Interest Points Based on Ecological Statistics

Neil Bruce* , Pierre Kornprobst†

Thème BIO — Systèmes biologiques
Projet Odysée

Rapport de recherche n° 6745 — December 2008 — 19 pages

Abstract: In this report, we consider whether statistical regularities in natural images might be exploited to provide an improved selection criterion for interest points. One approach that has been particularly influential in this domain, is the Harris corner detector. The impetus for the selection criterion for Harris corners, proposed in early work and which remains in use to this day, is based on an intuitive mathematical definition constrained by the need for computational parsimony. In this report, we revisit this selection criterion free of the computational constraints that existed 20 years ago, and also importantly, taking advantage of the regularities observed in natural image statistics. Based on the motivating factors of stability and richness of structure, a selection threshold for Harris corners is proposed that is optimal with respect to the structure observed in natural images. Following the protocol proposed by Mikolajczyk et al. [6] we demonstrate that the proposed approach produces interest points that are more stable across various image deformations and are more distinctive resulting in improved matching scores. Finally, the proposal may be shown to generalize to provide an improved selection criterion for other types of interest points. As a whole, the report affords an improved selection criterion for Harris corners which might foreseeably benefit any system that employs Harris corners as a constituent component, and additionally presents a general strategy for the selection of interest points based on any measure of local image structure.

Key-words: Harris corner detector, natural image statistics, selection criterion

* Neil.Bruce@sophia.inria.fr

† Pierre.Kornprobst@sophia.inria.fr

Le détecteur de Harris dans le monde réel: Un critère de sélection basé sur les statistiques des images naturelles

Résumé : Dans ce rapport, nous étudions si les régularités statistiques dans les images naturelles pourraient être exploitées pour fournir un critère de décision amélioré pour les points d'intérêt. Une approche particulièrement connue dans ce domaine est la méthode de détection de coins de Harris. Le choix du critère de décision proposé à l'origine, et qui est toujours utilisé actuellement, était basé sur une définition mathématique intuitive avec un souci d'efficacité numérique. Dans ce rapport, nous revisitons ce critère, libre de toute contrainte computationnelle telle qu'il existait il y a 20 ans, mais aussi en tirant profit des régularités observées dans les statistiques des images naturelles. Basé sur les facteurs de stabilité et de richesse des structures, un seuil de sélection pour les coins de Harris est proposé de façon à être optimal par rapport aux structures observées dans les images naturelles. En suivant le protocole proposé par Mikolajczyk et al. [6], nous montrons que l'approche proposée génère des points d'intérêt qui sont plus stables par rapport aux déformations de l'image et plus différenciés en améliorant les résultats de corrélations. Finalement, nous proposons l'idée de généraliser cette démarche pour améliorer les critères de sélection pour d'autres points d'intérêt. En résumé, ce rapport propose donc un nouveau critère de sélection pouvant être utilisé directement dans tout système utilisant la méthode de Harris, et il propose une stratégie générale pour la sélection de points d'intérêt basée sur une mesure de structure locale des images.

Mots-clés : Détecteur de coin de Harris, statistique des images naturelles, critère de décision

Contents

1	Introduction	4
2	Motivation and Approach	6
2.1	Motivation	6
2.2	Towards an Optimal Selection Criterion	6
3	Harris Corners and Natural Image Statistics	7
3.1	Methods	7
3.2	Results	8
3.3	Fitting	8
3.4	Scale	10
4	Harris corners and <i>real</i> corners	10
5	Quantitative Evaluation	12
5.1	The Harris Affine Operator	12
5.2	Evaluation Protocol	14
5.3	Test Images	14
5.4	Results	14
6	Discussion	17

1 Introduction

Interest operators have a long history in computer vision and remain a significant component in many machine vision systems, constituting an early feature extraction stage which typically guides higher level vision tasks. This involves the selection of a candidate set of points or regions possibly of varying scale and/or shape. This set of points/locations may then be used in object recognition, robot navigation, scene classification or a variety of other tasks. There are many proposals for the selection of interest points with the central criterion being invariance to deformations of the image and distinctiveness of local structure at chosen points; a consideration important for matching purposes. While many different definitions for the selection of interest points has been proposed, the most popular approach remains the Harris corner detector. The Harris corner detector was introduced two decades ago and now appears within hundreds of applications and has been cited more than 2500 times in published work at the time of writing.

The introduction of interest operators in the context of machine vision perhaps dates back to 1979 when an influential proposal for corner detection was put forth by Moravec [7]. Moravec's operator considers how similar a local region of the image is to nearby heavily overlapping regions, computing the sum of squared differences between the central region and regions in the local surround. That is, given an image patch centered at location i, j in an intensity image I this difference is given by $S(x, y) = \sum_i \sum_j (I(i, j) - I(i + x, j + y))^2$ and is computed for the neighbors in the horizontal and vertical directions as well as for the two diagonal directions. A corner is defined as a location that is locally maximal subject to $S(x, y)$. The intuition behind this procedure is that for homogeneous (flat) regions, this difference will be very small with the difference becoming greater for regions containing edges and even larger for regions containing corners.

Harris and Stephens refined this idea by considering directional derivatives in lieu of shifted patches to produce a more robust corner detector, with invariance to rotation [2]. Specifically, $I(i + x, j + y)$ becomes $I(i, j) + I_x(i, j)x + I_y(i, j)y$ subject to a truncated first order Taylor series expansion where I_x and I_y are partial derivatives in x and y respectively. A substitution of this term into $S(x, y)$ and inclusion of a weighting parameter $w(i, j)$ yields the expression $S(x, y) = \sum_i \sum_j w(i, j)(I_x(i, j)x + I_y(i, j)y)^2 = (x \ y) A (x \ y)^T$ where

$$A = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix}$$

where angle brackets denote summation over i, j subject to the weighting function $w(i, j)$. The function w is defined here as a gaussian function of variance σ which defines the scale of the analysis. The matrix A describes the local intensity structure of the neighborhood centered around i, j . The judgement of whether a pixel location corresponds to a corner is based on the eigenvalues λ_1 and λ_2 of the matrix A . Specifically, when $\lambda_1 \approx 0$ and $\lambda_2 \approx 0$ a flat region is found, when λ_1 is a large positive value and $\lambda_2 \approx 0$ an edge is found. When λ_1 and λ_2 are both large positive values, a corner is present. Following this intuition, it has been proposed that interest points be selected according to the locations for which

$$\lambda_1 \lambda_2 - \kappa(\lambda_1 + \lambda_2)^2 = \det(A) - \kappa \text{tr}^2(A)$$

is large and a local maximum with κ a constant. This produces a strong value when both λ_1 and λ_2 are large but penalizes situations where one of these values is much larger than the other indicative of an edge situation. There are two specific concerns that one might raise considering the form of this expression:

1. The specific form of this expression is one of *intuitive* quantitative reasoning (based on the determinant and trace of A) but significantly limits the shape of the decision boundary for choosing candidate Harris corners. This specific form is one based on computational parsimony since the determination of the eigenvalues of A requires the computation of a square root. While this operation may have been sufficiently cumbersome (in a computational sense) in 1988 to warrant a selection criterion that avoids computing square roots, this is much less of a concern with modern computing hardware. The concern related to computation of square roots arises from the fact that an eigendecomposition for the 2x2 case is given by the quadratic formula with

$$\lambda_{1,2} = \frac{1}{2}(\text{tr}(A) \pm \sqrt{\text{tr}^2(A) - 4 \det(A)}) \quad (1)$$

2. It is unclear what the specific tradeoff between pairs of intermediate magnitude eigenvalues and one large eigenvalue paired with one smaller eigenvalue (given by the parameter κ should look like. This is apparent from the fact that there is no universal consensus on the value of κ which from a system performance perspective will most likely be application dependent. Anecdotal observations suggest a range of 0.04 to 0.15 as appropriate choices.

While the definition of local structure that constitutes Harris corners is sensible, it is arguable that the form of the selection boundary deserves further consideration. Moreover, the original form proposed in 1988 remains in common use and remains the most common means of selecting interest points in the machine vision literature.

This brings us to the central motivation of this paper: To revisit Harris corners from the perspective of the decision criterion employed and in doing so, to provide a sensible principled decision boundary for selecting Harris corners on the basis of principles that may generalize to any interest operator for which there exists a selection criterion based on local structure.

Many other structural definitions for Harris corners have been proposed but only very few that consider the impetus for the selection criterion based on the eigenvalues of the autocorrelation matrix. This is surprising in light of the fact that the basic structure of Harris detector seems to imply a detector that is highly robust to changes in rotation and scale [10]. Existing efforts that focus on the decision boundary associated with Harris corners are limited in that are either *ad hoc*, based on heuristics [8, 9] or focus on the approximation of $I(i+x, j+y)$ [9].

There is heretofore no existing work that attempts to construct a definition for the selection of Harris corners based on first principles and motivated by the criteria of invariance and distinctiveness. In this paper, we put forth such a proposal which provides a definition for the decision boundary based on these motivating principles through observation of the structure of natural images as it pertains to Harris corners. Additionally, this definition exploits the statistics of the natural world in order to best satisfy these criteria.

The format of this paper is as follows: Section 2 provides a more detailed description of the precise set of criteria that might motivate the decision boundary for selecting Harris corners. This

also includes a statement of the central premise of our proposal in light of these criteria. In section 3 the statistics of natural images are considered as they relate to Harris corners through observation of a large set of natural images. Commentary and modeling pertaining to the observed statistics is presented. Section 4 provides some qualitative results on natural images containing geometric forms and discussion is included concerning the relationship between corners chosen by the Harris algorithm and *real* corners appearing in the natural world. Section 5 presents quantitative evidence in favor of the proposal through consideration of the stability of point selection under various natural image deformations including combined rotation and zoom, and changes in viewpoint revealing advantages of the proposed selection criterion as compared with the *traditional* approach. Finally, in section 6 we discuss more general issues pertaining to the proposal at hand and present many possible fruitful avenues for further research.

2 Motivation and Approach

2.1 Motivation

As mentioned, interest point detection typically involves the extraction of features with a specific application domain in mind. Typical applications include at least image matching, tracking, panorama stitching, 3D-modeling and object recognition among others. The central criteria for candidate interest points (and the evaluation criterion associated with such points) are typically:

1. **Repeatability:** For purposes such as matching across viewpoints, tracking, and stitching, an important property is that the set of discrete interest points selected by an algorithm for one view corresponds to the same points in the world drawn from a different viewpoint, or subject to some transformation/deformation of the image.
2. **Distinctiveness:** Any process that relies on matching candidate points across a change in viewpoint, time or other deformation of the image requires that the points chosen correspond to regions with distinct structure in order for points from say, one view point *to be matched with* those from another.
3. **Geometry:** For a task such as 3D modeling, one may strongly desire that the interest points correspond strictly to some geometric construct. For example, the most desirable property may be that a corner detector produces well localized corners that correspond to real corners within the image, ignoring repeatability and distinctiveness.

In light of these motivating factors, it is unclear to what extent the selection boundary for Harris corners proposed in the original work [2] is suited to these conditions.

2.2 Towards an Optimal Selection Criterion

One sensible manner in which to approach the problem is in considering the sort of structure that is observed within natural images. Specifically, it may be sensible to consider the observation likelihood of the two eigenvalues $p(\lambda_1, \lambda_2)$ as a means of determining a decision threshold. The impetus for this choice with respect to the three motivating factors is as follows:

1. **Repeatability:** For any given interest point h , one has corresponding values λ_1 and λ_2 . A corresponding point in a second image (e.g. the first image subject to a slight change in zoom/rotation/viewpoint) will have the effect of perturbing the eigenvalues in question (corresponding to the same point) by some small values δ_1, δ_2 . It may be the case that while (λ_1, λ_2) lies within the selection boundary and is deemed an interest point, $(\lambda_1 + \delta_1, \lambda_2 + \delta_2)$ may lie outside of the selection boundary. In terms of an overall repeatability score (as the number of points in image 1 that have corresponding points in image 2) it is desirable to maximize the distance (on average) between interest points chosen, and the selection boundary as this reduces the likelihood that such a perturbation will result in a point moving from inside to outside of the decision boundary (i.e. the tolerance for δ_1 and δ_2 is higher). It suffices to choose interest points based on some threshold T such that $p(\lambda_1, \lambda_2) > T$ to satisfy this condition.
2. **Distinctiveness:** It is also evident, that in choosing interest points as an inverse function of the likelihood of local structure parameters, one also has a desirable property with respect to distinctiveness. In fact, this is the very definition of distinctiveness: Points whose local structure is observed least frequently are chosen first, followed by those that appear with increasing frequency. It is interesting to note that this is precisely the criterion employed to measure the distinctiveness of interest points that appears in [10].
3. **Geometry:** With respect to local geometry, the relationship between the likelihood of the local structure coefficients and specific real world constructs is less evident. That being said, on an intuitive level it is expected that local structure corresponding to corners should appear much less frequently than that corresponding to edges as observing ones surroundings at any given time may well confirm. For this reason, it is reasonable to assume that one might also arrive at a selection criterion that is favorable for applications involving 3D reconstruction, or determination of geometric form.

As a whole, the choice of a decision boundary based on the reciprocal likelihood of observed local structure parameters (e.g. λ_1 and λ_2 in the case of Harris corners) seems a sensible strategy for the choice of a decision boundary with optimality in some sense with respect to repeatability and distinctiveness, and also at an intuitive level in its correspondence to image geometry.

3 Harris Corners and Natural Image Statistics

In this section, we describe methods and results associated with deriving an estimate of $p(\lambda_1, \lambda_2)$ in the context of all natural images.

3.1 Methods

2100 images were drawn from the Corel stock photo database, consisting of indoor and outdoor scenes with photographs taken at several venues around the world. The central 1200x800 pixel portion of the images was cropped and the eigenvalues of the Harris matrix computed for each pixel location across the 2100 images. A histogram density estimate was constructed with a bin width of 0.0061 by 0.0061.

3.2 Results

Figure 1 depicts the log of $p(\lambda_1, \lambda_2)$ for $\sigma = 2$ as determined from the 2100 images in the test set with values reflected across the $y = x$ axis since $\lambda_1 > \lambda_2$ as given by equation 1. Additionally, the traditional Harris criterion boundary for several different thresholds for $\kappa = 0.15$ shown in red and for $\kappa = 0.04$ shown in navy are presented.

Notably, there is a significant difference between the shape of the surface produced by considering $p(\lambda_1, \lambda_2)$ relative to the Harris decision boundaries. Also of note, is the fact that the lowest probabilities correspond to the $\lambda_1 = \lambda_2$ case, which agrees with our previous intuition concerning the manner in which geometry relates to observed statistics. That is, the equality case for the eigenvalues seems to correspond most closely with the definition of a *perfect* corner and corresponds to the least likely structures in the statistics.

A second salient observation may be made concerning the distribution itself: As the dropoff in $p(\lambda_1, \lambda_2)$ is much steeper in the direction of increasing λ_2 than in the direction of λ_1 , and λ_1 is at least as large as λ_2 this should make λ_2 a very good predictor of $p(\lambda_1, \lambda_2)$ (recall that points only exist in the upper left sector of figure 1). It is interesting to note that selection of interest points based on $\min(\lambda_1, \lambda_2)$ appears in one effort with the motivation for this choice based on choosing good features for the purpose of tracking [11]. The motivation for this choice is that in practice this produces points for which the autocorrelation matrix is above the noise level and is well-conditioned. This particular measure is also mentioned in the context of a study involving the comparison of interest points [10], but this particular choice is not included in the comparison. There also exists an additional effort by Tsai and Su [12] that considers the predictive capacity of the smaller eigenvalue to locate corners based on generated analytic structures and the authors conclude that the smaller of the two eigenvalues is a very good predictor of corners. There is however no mention of the Harris measure in this work and thus no comparison. The minimum eigenvalue approach seems a good candidate from a theoretical standpoint as it verifies a set of important axiomatic properties [3] that other measures including the Harris criterion fail to satisfy. Here, we also demonstrate that the likelihood estimate provides a strong case for the minimum eigenvalue as a measure of corneredness in support of the suggestions made in the aforementioned studies.

3.3 Fitting

One might select interest points on the basis of a direct lookup on the likelihood of λ_1, λ_2 given by the probability density function, choosing points in increasing order of $p(\lambda_1, \lambda_2)$. Within a histogram density estimate, the nature of discrete computation requires quantization in the form of bin size. This quantization has the possibility of causing various point locations within the image to be assigned identical scores complicating the task of choosing a candidate point set. Furthermore, using a sufficiently fine quantization (as used in figure 1) to diminish the likelihood of this occurring, even for an estimate based on more than 2 trillion samples as we have performed, one still encounters λ_1, λ_2 combinations for several locations within an image that have only been observed a small number of times (including 0), making the order of point selection ambiguous. This is due to the exponential dropoff in observations with an increase in λ_1 or λ_2 .

A solution to this is to fit a function to $p(\lambda_1, \lambda_2)$ affording a continuous representation of $p(\lambda_1, \lambda_2)$.

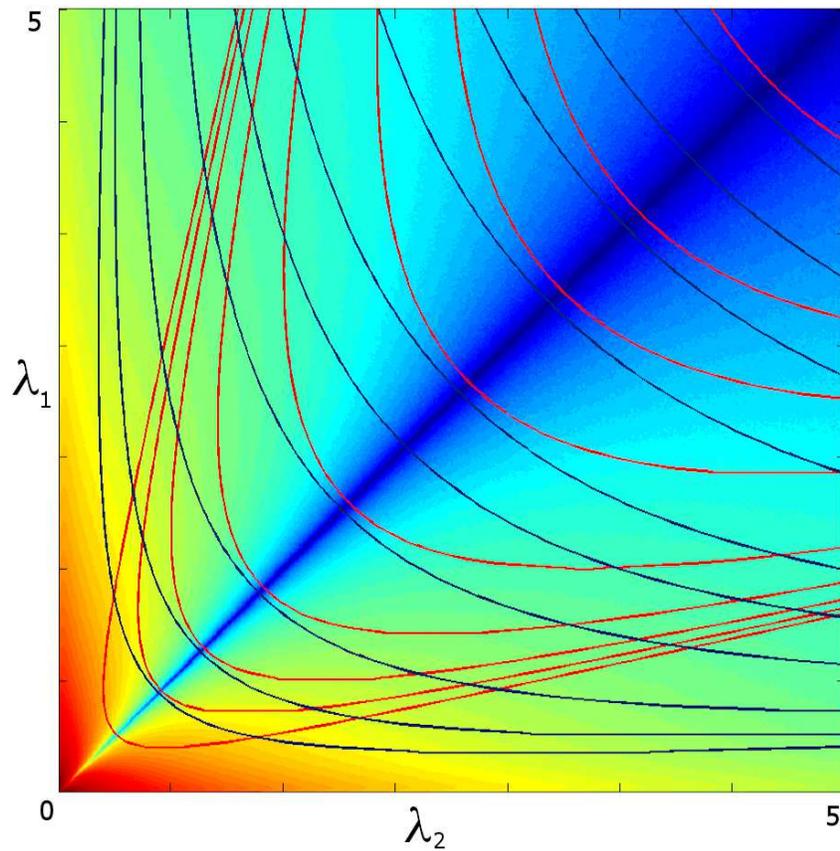


Figure 1: The log likelihood of the various possible combinations of λ_1, λ_2 in the context of a large sample of natural images, with the Harris criterion for several thresholds corresponding to $\kappa = 0.04$ (navy) and $\kappa = 0.15$ (red) superimposed. Note that there exists a significant difference in the shape of the threshold given by the traditional Harris criterion relative to the value of $p(\lambda_1, \lambda_2)$ in the probability density function on the pair of eigenvalues of the autocorrelation matrix.

It should be stressed that this relies on the reasonable assumption that the distribution observed on the basis of smaller eigenvalues remains the same for larger eigenvalues.

There are many possible assumptions that one might make about the form of this function. Figure 1 suggests that a polynomial fit to $\log(p(\lambda_1, \lambda_2))$ might be adequate. We have performed a minimum least squares fit to the learned $\log(p(\lambda_1, \lambda_2))$ distribution based on a polynomial of order n for $n \in \{2, 3, 4, 5\}$. Corresponding functions are depicted in figure 2. It is interesting to note that a high order polynomial fit is required to capture the *long tailed* nature of the distribution suggesting a simpler form or better fit might be achieved via a non-polynomial fit of some form. This however is tangential to the point of this paper but may present an interesting avenue for future consideration and the polynomial fit we have considered appears to be sufficient for the purposes of this study with the $n = 5$ case producing a mean squared error of less than 10^{-15} . The order 5 polynomial is given by $P_{ij} =$

$$\begin{bmatrix} -0.0042 & 0.1014 & -0.9998 & 5.0371 & -13.0068 & 13.7893 \\ 0.0145 & -0.3589 & 3.6914 & -20.0508 & 57.0538 & -67.4095 \\ -0.0078 & 0.1991 & -2.2568 & 19.2248 & -73.3100 & 110.1198 \\ -0.0156 & 0.3808 & -2.9872 & 6.2377 & 18.1182 & -69.2591 \\ 0.0156 & -0.3881 & 3.4473 & -12.7242 & 14.4629 & 15.2591 \\ -0.0052 & 0.1298 & -1.2093 & 5.2478 & -10.9142 & 10.3522 \end{bmatrix}$$

where the coefficient appearing at P_{ij} corresponds to the term $x^i y^j$ with $i = j = 5$ corresponding to the top-left element, and with $x = \lambda_1 - \lambda_2$ and $y = \lambda_1 + \lambda_2$.

3.4 Scale

While the original Harris operator is rotationally invariant, many recent efforts consider the application of the Harris operator at several spatial scales suggesting the need to consider the distribution of $p(\lambda_1, \lambda_2)$ for different choices of σ . Owing to the scale invariance property of natural images [1], one might expect that the distribution of $p(\lambda_1, \lambda_2)$ may share this same property. From a qualitative perspective, the distributions produced by the eigenvalues of the autocorrelation matrix are very similar across scale. That said, for large changes in scale, the exponential dropoff in the direction of $\lambda_1 + \lambda_2$ seems to diminish slightly with increasing scale. We also observed that the ranking of points for a given image according to $p(\lambda_1, \lambda_2)$ remains largely the same regardless of the scale at which the distribution is learned. For certain applications, in particular those for which comparison of likelihoods across different scales is important, this *might* be a significant factor. For the purposes of this study the results are computed based on the distribution learned from the $\sigma^2 = 4$ condition shown in figure 1 and this should not have any significant impact on the analysis or conclusions.

4 Harris corners and *real* corners

In instances where the aim is to perform some form of geometric modeling of the scene, a sensible goal is to ensure that the selected corners conform to corners of real physical entities. The following provides an example of the operation of the original Harris criterion as compared with the likelihood based criterion as applied to some natural images with strong geometric structure. The top image appearing in figure 3 is a common test image in the domain of interest point detection, and the second similar image is chosen for its geometric structure. In each case, a candidate set of points for

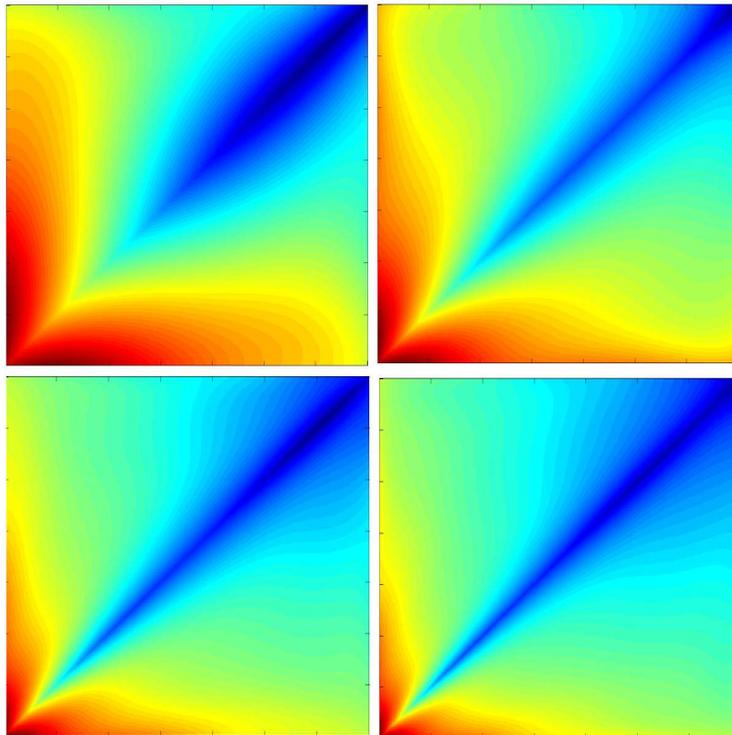


Figure 2: An example of polynomial fits (top left, top right, bottom left, bottom right) of order $n \in \{2, 3, 4, 5\}$

the traditional Harris threshold with $\kappa = 0.04$ is shown (left) and for the likelihood based selection threshold (right) choosing points that score highest in order to equalize the number of points selected. It appears that the likelihood based selection criterion results in more corners corresponding to the physical corners of objects and additionally, better localization in some instances. Notice for example in the top figure that several of the corners of the hexahedron are correctly selected by the likelihood based measure but are missed by the Harris criterion. The points located at the back of the hexahedron (e.g. the one marked with the asterisk (*)) are symptomatic of one issue with the standard Harris criterion: These points consist of one eigenvalue that is very large combined with a smaller eigenvalue of intermediate size. The strength of the larger eigenvalue prevents selection of the corner even for very small κ values. Also note the mislocalization of some of the corners on the blocks.

The second image points out an additional issue pertaining to the standard Harris criterion. The eigenvalues may differ by orders of magnitude such that the order of strength of this measure is frequently corner \leq flat region \leq edge. This is evident in the points selected in central regions of the geometric structures. In the interest of forming a geometric representation of the scene, it is sensible to enforce an ordering of corner \leq edge \leq flat region so that the chosen points that are not corners nevertheless correspond to the geometry of the figure.

5 Quantitative Evaluation

On the basis of the discussion appearing in section 2 it is sensible to provide a simple demonstration that the proposal of likelihood based interest point selection yields better performance in terms of repeatability in a quantitative sense. As the proposal at hand provides the possibility of improved performance for any application that employs Harris corners as a constituent component, there are many choices that one might make to provide some quantitative analysis consistent with the proposal.

For this purpose, we have chosen an evaluation paradigm for the selection of interest points that is in common use and evaluates repeatability and matching scores for affine regions selected on the basis of the Harris corner criterion with repeatability and matching scores considered within this evaluation paradigm. The extension of the Harris corner detector to the selection of affine regions and the specifics of the evaluation follow.

5.1 The Harris Affine Operator

Mikolajczyk and Schmid propose a modification of the Harris detector to make it suitable for the selection of affine regions [5]. The specific details are beyond the scope of the current work, but in short the Harris operator is applied at several spatial scales and maxima are selected according to the traditional measure of Harris corneredness. For points that are above threshold, the characteristic scale is identified and regions normalized based on affine shape adaptation. Scale parameters and localization are then estimated via an iterative algorithm. For the specific details, the reader is encouraged to consult [5].

In the following study we have computed affine regions according to this strategy, with the best 500 candidate points chosen for scales of $\sigma^2 \in 2, 4, 8, 16$. Following affine shape adaptation and

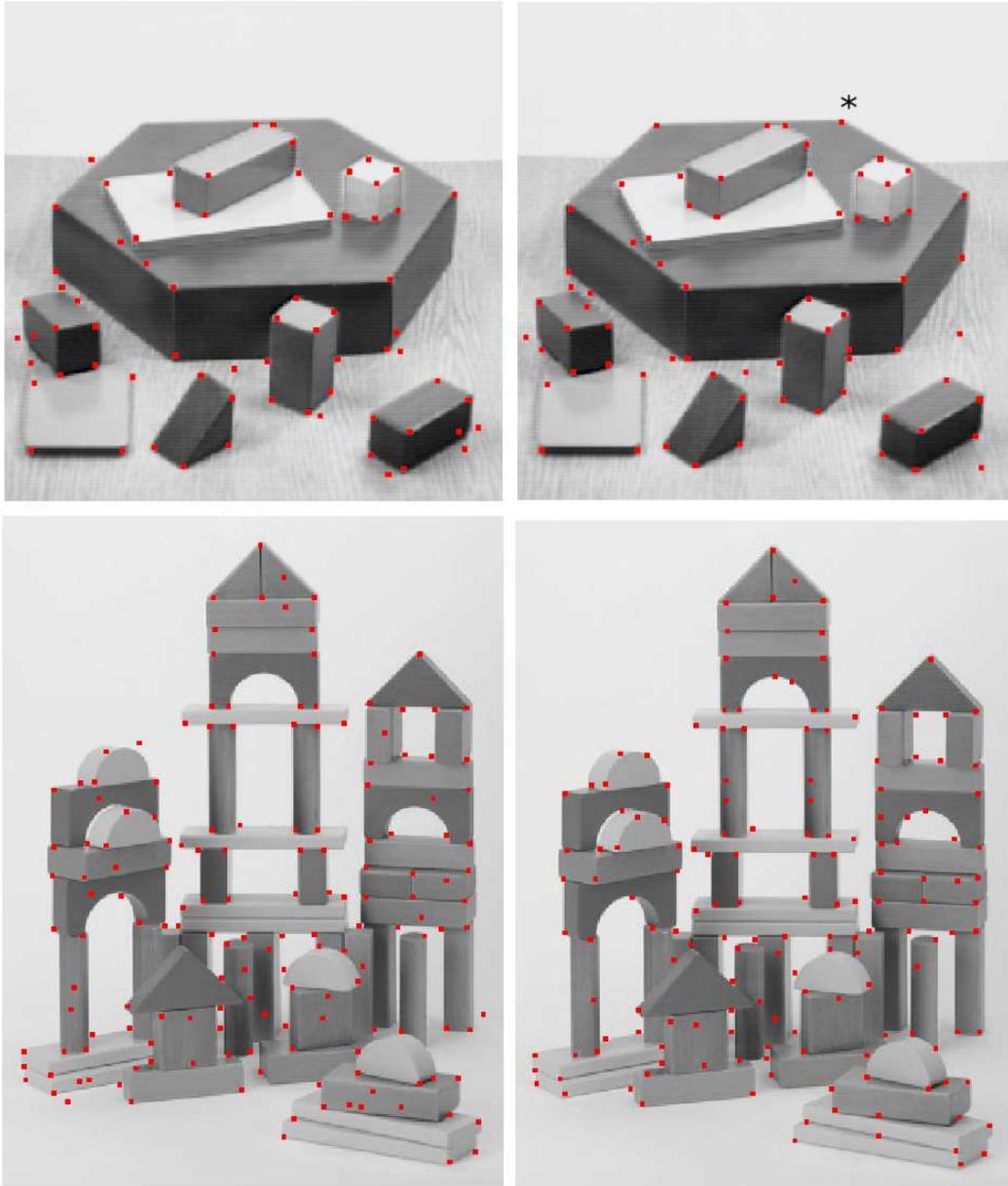


Figure 3: An example of corners selected by the original Harris criterion (left, with $\kappa = 0.04$) and by the likelihood based criterion (right) for a *classic* test image from the corner detection literature (a), and for a similar example.

rejection of points that fail to converge this typically leaves approximately 1800 candidate regions in each image for which the number in the corresponding region between any pair of images varies. We have employed the implementation of affine shape adaptation of Laptev [4].

5.2 Evaluation Protocol

In the evaluation protocol put forth by Mikolajczyk, Schmid et al. [6] overlap between two regions is defined according to the measure $1 - \frac{R_{\mu_a} \cap R_{(H^T \mu_b H)}}{(R_{\mu_a} \cup R_{(H^T \mu_b H)})}$ with R_{μ} corresponding to the elliptical region defined by $x^T \mu x = 1$ and with H the homography relating the two images from which R_{μ_a} and R_{μ_b} are drawn.

A repeatability score may then be assigned to any pair of images based on the number of regions that exhibit an overlap score greater than some value ε relative to maximum number of possible overlapping regions. Each image contains some number of points within the overlapping region defined by the homography and the maximum number of possible overlapping regions equals the minimum of the number of points within this overlapping region for the two images. In the results presented in [6] results for an overlap of 40% are employed and this is also the threshold for overlap in the results reported in this section.

5.3 Test Images

The test set employed is drawn from the images considered in [6] and a subset of these images is depicted in figure 4. Images consist of two natural scenes, one subjected to combined rotation and scale and one subjected to a change in viewpoint, and two textured images, with the same two types of deformations as the natural scenes. It is perhaps worth stating once again, that one might consider any evaluation paradigm for the purposes of observing the stability of points selected by the algorithm but this particular choice is natural given the popularity of this evaluation paradigm and the availability of source code for affine region selection and evaluation of repeatability.

5.4 Results

Figure 5 demonstrates the repeatability scores associated with increasing degrees of combined rotation/zoom and change of viewpoint. The results confirm a few points made earlier. First, note that the repeatability (and set of candidate regions) is similar for the proposed likelihood based approach and the minimum eigenvalue approach confirming that the minimum of the eigenvalues is a good estimate of the likelihood of the eigenvalue pair. The likelihood and minimum eigenvalue demonstrate clearly better repeatability scores for both the natural image and texture subjected to combined rotation and zoom. In the case of a viewpoint change, the likelihood and minimum eigenvalue selection is clearly better for smaller viewpoint changes, but this advantage diminishes with increasing viewpoint angle (and even becomes a deficit for the graffiti image for the most extreme viewpoint angles). There are a few possible explanations for this behavior. One explanation for this behavior pertains to the nature of the training set. Although the Corel stock photo database consists of a large variety of natural images, these images are largely composed images and thus confined



Figure 4: Test images from the study of Mikolajczyk et al. the images are subjected to natural deformations including changes in scale and rotation as in the bark and boat examples (top row, second row) and changes of viewpoint (third row, fourth row).

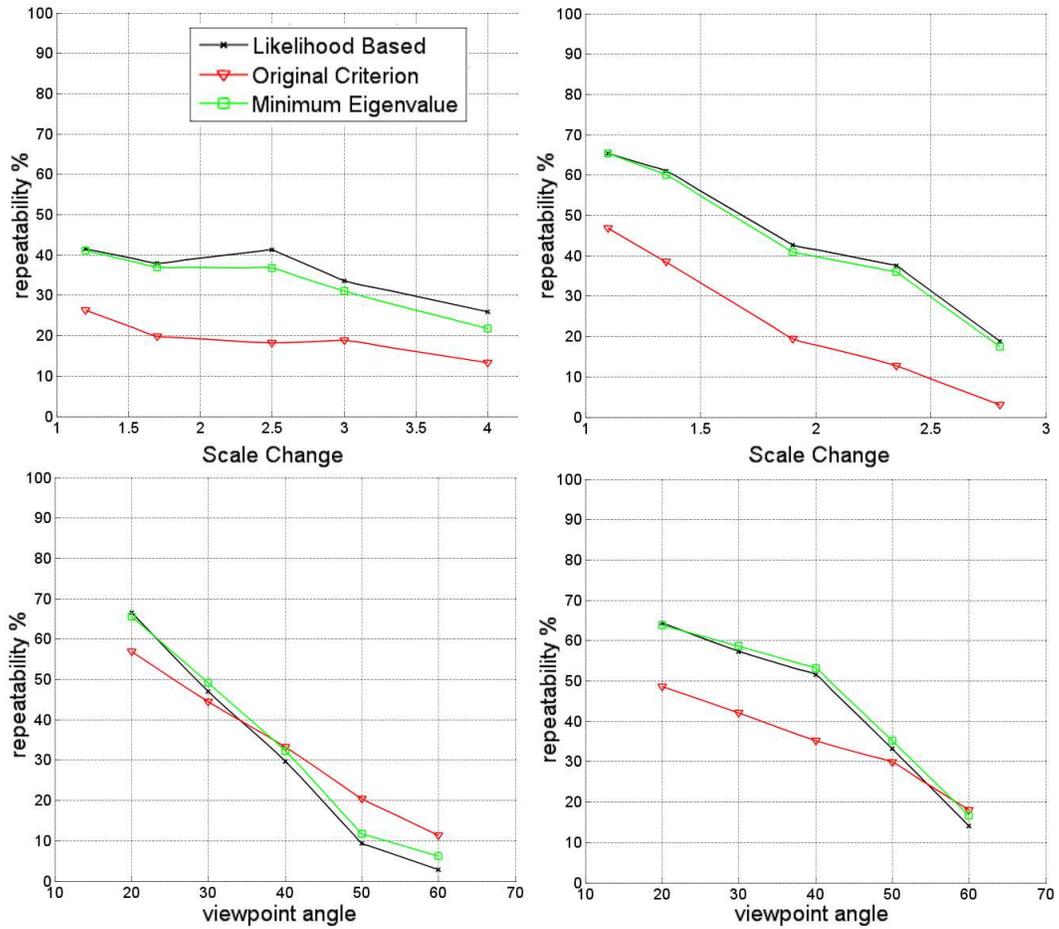


Figure 5: Repeatability scores across increasingly severe deformations for the bark example (top left), the boat example (top right), the graffiti example (bottom left) and the wall example (bottom right).

to frontoparallel views. Thus it is possible that for a very severe deformation such as the rightmost graffiti image shown in figure 4 that the statistics of the image are sufficiently different from the training set that the previously selected corners are deformed in a manner that they are no longer improbable according to the training set. In the case of a mobile robot performing navigation within a certain environment and roaming in order to collect representative views, this would presumably not be an issue. A second possibility is that the effect of any given deformation on the values of λ_1 and λ_2 is systematic in nature. The proposal should imply optimality under an isotropic perturbation of λ_1 and λ_2 associated with some image deformation. That said, a severe change in viewpoint angle may perturb λ_1 and λ_2 in a specific manner such that the original Harris criterion happens to be less sensitive to said perturbation. In this case, it is less important how close points are to the decision boundary and more important is, in which direction these points tend to move (in the λ_1, λ_2 plane). This also suggests an interesting possible avenue for future research, that being the effect that specific types of deformations have on the values of λ_1 and λ_2 providing a stronger optimality criterion for the decision boundary.

6 Discussion

There are a variety of interesting points that emerge from the subject matter of this paper. We have presented an analysis of the selection of interest points based on the Harris corner detector revisiting a central aspect of the algorithm motivated by the specific purposes for which interest points are employed. Motivated by these design criterion, we propose a selection criterion that is optimal from the perspective of natural image statistics. It should perhaps be emphasized that our intention is not to present a general closed form formula for choosing Harris corners for any and every purpose (although the results given may satisfy this purpose to some extent). It is possible, or even likely that for a purpose such as localization in mobile robot navigation, that one may do even better in constructing an estimate of $p(\lambda_1, \lambda_2)$ that accounts for the specific statistics of the environment in which the robot is navigating. Furthermore, more detailed analysis may produce a better fit to the distribution in question. It should be stressed that this is a case in which the message is more important than the medium: that a likelihood based selection criterion that considers the relevant statistics is a natural way of choosing interest points. The methods put forth in this paper provide a general guideline for application specific Harris corner selection and the specific choice of a fit to the probability density is based on simplicity of exposition and may benefit from further analysis. In short, the aim is to demonstrate that a likelihood based selection criterion that exploits natural image statistics may produce more stable and distinctive points of interest. That said, conclusions and points of interest emergent from this study are as follows:

- A likelihood based selection criterion that considers the structure observed in the statistics of natural images is a sensible means of choosing points of interest on the basis of qualitative and quantitative evidence
- Although we have considered Harris corners, the motivation and proposed principle is not specific to Harris corners and may be equally applied to any strategy for interest point selection for which there is a choice of selection boundary

- The strategy of likelihood based selection presents the possibility of producing a selection criterion that is environment and/or task specific by choosing a particular training set (e.g. images from a roaming robot, or an exhaustive systematic sample of images that represent a specific deformation such as a change of viewpoint)
- The minimum of the two eigenvalues of the autocorrelation matrix appears to provide a good approximation to the likelihood of the eigenvalue pair. The use of a minimum eigenvalue selection criterion does appear in a few places in the literature, but a thorough comparison of this selection criterion with the traditional Harris measure is warranted
- The likelihood based criterion selects more corners that correspond to real corners of objects making it amenable to the representation of image geometry and the apparent failure of the traditional measure seems to be in penalizing locations for which two strong eigenvalues are present, one of which is of exceptional magnitude or disallowing instances in which both eigenvalues are of smaller magnitude but nearly identical
- The results presented in section 5 suggest an additional avenue for future consideration. Specifically, the study of the manner in which some systematic deformation alters the parameters (the two eigenvalues in this case) on which the selection criterion is based. Section 5 also suggests that the proposed methodology may be sensitive to the data set employed and the precise contribution of these two elements is unclear.
- Even in the presence of anisotropic perturbation of eigenvalues based on an image deformation, small deformations of the image should result in optimal repeatability as evidenced by the data presented. It is worth noting also that many of these tasks involve deformations that are far less severe than say, that between the leftmost and rightmost graffiti images shown in figure 4 and thus performance improvements may also be observed for many typical tasks without the need for a *stronger* model.

As a whole, we have presented a novel principled strategy for the selection of Harris corners motivated by structure observed in natural image statistics. The analysis presented raises many interesting questions pertaining to the selection of Harris corners, and the nature of interest operators in general. As a whole, the proposal presents the possibility of improved performance for any system that employs Harris corners as a constituent component and presents many fruitful directions for further research.

Acknowledgments

This work was partially supported by EC ICT PROJECT No. 215866 (SEARISE)

References

- [1] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Am A*, 4(12):2379–2394, December 1987.
- [2] C. Harris and M. Stephens. A combined corner and edge detection. In *Proceedings of The Fourth Alvey Vision Conference*, pages 147–151, 1988.

- [3] C. S. Kenney, M. Zuliani, and B. S. Manjunath. An axiomatic approach to corner detection. In *2005 Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, pages 191–197, 2005.
- [4] I. Laptev. www.nada.kth.se/~laptev/code.html.
- [5] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. In *In Proc. European Conf. Computer Vision*, pages 128–142. Springer Verlag, 2002.
- [6] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *Int. J. Comput. Vision*, 65(1-2):43–72, 2005.
- [7] H. Moravec. Obstacle avoidance and navigation in the real world by a seeing robot rover. In *CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University*. 1980.
- [8] N. Overgaard. On a modification to the harris corner detector. pages 6–7. Symposium Svenska Sällskapet för Bildanalys, Stockholm, 2003.
- [9] S.-C. Pei and J.-J. Ding. Improved harris’ algorithm for corner and edge detections. In *ICIP (3)*, pages 57–60, 2007.
- [10] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172, 2000.
- [11] J. Shi and C. Tomasi. Good features to track. In *Conf. on Computer Vision and Pattern Recognition*, pages 593–600, 1994.
- [12] D.-M. Tsai, H.-T. Hou, and H.-J. Su. Boundary-based corner detection using eigenvalues of covariance matrices. *Pattern Recogn. Lett.*, 20(1):31–40, 1999.



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399