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

Neil D.B. Bruce and Pierre Kornprobst
INRIA, 2004 route des Lucioles, BP 93, 06902, Sophia Antipolis, France
http://www-sop.inria.fr/members/Neil.Bruce/

Abstract

In this paper, 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 paper, 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 based on a definition of optimality with respect to the structure observed in natural images. As a whole, the paper affords considerable insight into why existing approaches for selecting interest points work, and also their shortcomings. We also demonstrate how a proposal that is inspired by the properties of natural image statistics might be applied to overcome these shortcomings.

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 have 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 [3]. 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 $S(x,y) = \sum_i \sum_j w(i,j)(I_x(i,j)x + I_y(i,j)y)^2 = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x & y \end{pmatrix}^T$ where

$$A = \begin{bmatrix} < I_x^2 > & < I_x I_y > \\ < I_x I_y > & < I_y^2 > \end{bmatrix}$$

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

978-1-4244-3991-1/09/$25.00 ©2009 IEEE

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 $\lambda_1\lambda_2 - \kappa(\lambda_1 + \lambda_2)^2 = \det(A) - \kappa \mathrm{tr}^2(A)$ 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
$$\lambda_{1,2} = \frac{1}{2}(\mathrm{tr}(A) \pm \sqrt{\mathrm{tr}^2(A) - 4\det(A)})$$
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 [11]. Existing efforts that focus on the decision boundary associated with Harris corners are limited in that are they 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 a principled approach 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 existing algorithms, 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 and examines systematically the effect of large 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

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 in 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 viewpoint *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 or more generally, a set of points that conforms to the geometric structure of a scene.

In light of these motivating factors, it is unclear to what extent the selection boundary for Harris corners proposed in the original work [3] is suited to these conditions.

2.1. 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 change in zoom/rotation/viewpoint) will have the effect of perturbing the eigenvalues in question (corresponding to the same point) by some 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 (or the converse, 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 [11].
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 and the same made be said with regards to edges when compared with regions of low activity. 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. 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 yielding a total of approximately $2 * 10^9$ observations. A histogram density estimate was constructed with a bin width of 0.5 by 0.5.

3.1. 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. One salient observation that may be made concerning the distribution itself is as follows: 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 $p(\lambda_1, \lambda_2)$. It is interesting to

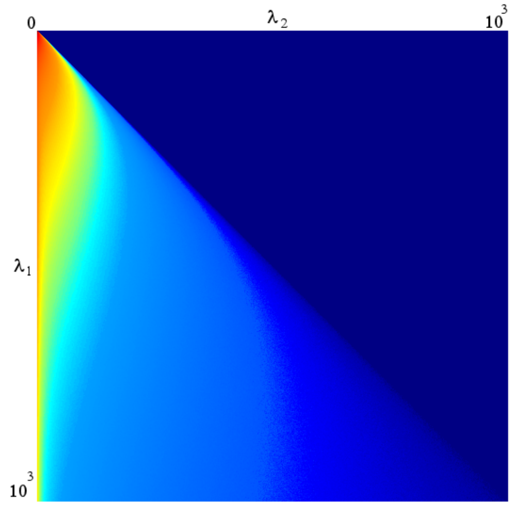

Figure 1. The log likelihood of the various possible combinations of λ_1, λ_2 in the context of a large sample of natural images. Note the concentration of points around the origin and the asymptotic distribution of density along the λ_1 axis. For larger values of λ_1 the dropoff along the λ_2 axis is much steeper than for smaller values of λ_1.

note that selection of interest points based on $min(\lambda_1, \lambda_2)$ appears in the proposal of Shi and Tomasi with the motivation for this choice based on choosing good features for the purpose of tracking [12]. 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. On the basis of this observation, comparisons made in the evaluation section also include the Shi criterion in particular because comprehensive evaluations of interest point selection [11] have not included this particular metric in their evaluations. The minimum eigenvalue approach seems a good candidate from a theoretical standpoint as it verifies a set of important axiomatic properties [4] 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.

3.2. 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 very few (including 0) times. The sparsity involved due to the exponential dropoff in observations with an increase in λ_1 or λ_2 can therefore make the order of point selection ambiguous in particular paired with non-maximal suppression. A solution to this is to fit a function to $p(\lambda_1, \lambda_2)$ affording a continuous representation of $p(\lambda_1, \lambda_2)$.

There are many possible assumptions that one might make about the form of this function. Early experiments with high order polynomial fits yielded satisfactory behavior, but the nature of these fits resulted in significant sensitivity to the range and distribution of eigenvalues present. In particular, such fits often converge to a state that is either overly biased to the form near the origin, or in the case that only a higher range of λ_1 values are considered the polynomial becomes relatively ambivalent to λ_1 and does not represent the structure of the distribution close to the origin. A further commentary on this appears in section 5.1. As $p(\lambda_1, \lambda_2)$ decreases exponentially in λ_1 and λ_2, for the purposes of exposition it may suffice to consider a least squares fit to $log(p(\lambda_1, \lambda_2))$ based on the simple form $C - \alpha\lambda_1^\beta\lambda_2^\gamma$, to afford some sense of the tradeoff between λ_1 and λ_2. This will in practice provide at a minimum, the opportunity to observe the tradeoff between λ_1 and λ_2 explicitly in a manner consistent with the observed statistics. A least squares fit to $log(p(\lambda_1, \lambda_2))$ of the form $C - \alpha\lambda_1^\beta\lambda_2^\gamma$ yields values of $\alpha = 0.4$, $\beta = 0.197$ and $\gamma = 0.322$. This affords as a classifier the threshold criterion $T > \lambda_1^\alpha\lambda_2^\beta$. This fit along with the traditional criteria are pictured in figure 2. Note that that the resultant fit has some similarities to both of the other classic measures shown, but importantly does not penalize large values of λ_1 and exercises some tradeoff between λ_1 and λ_2 especially for low values of λ_1.

3.3. Scale

While the original Harris operator is rotationally invariant, many recent efforts consider the multi-scale applications of the Harris operator, 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 [2], one might expect that the distribution of $p(\lambda_1, \lambda_2)$ may share this same property. From a qualitative perspective, the shape of distributions produced by the eigenvalues of the autocorrelation matrix are very similar across scale. That said, for large changes in scale, the range and distribution of observed eigenvalues does vary slightly. For certain applications, for which comparison of likelihoods across different scales is important, this *may* be a significant factor. For the purposes of this study the results are computed based on

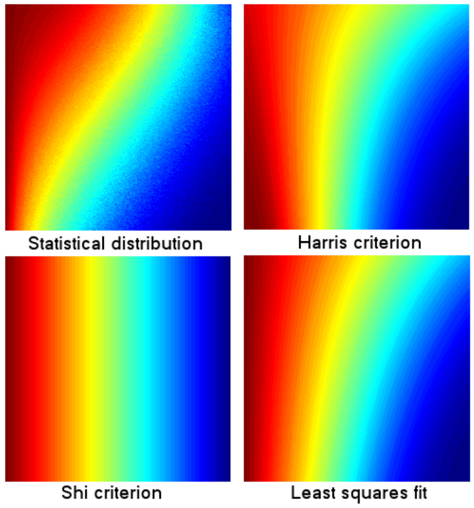

Figure 2. A surface representation of the decision boundaries for various choices of threshold for eigenvalues in a representative range given by $\lambda_1 \in [300, 800]$ and $\lambda_2 \in [1, 200]$ is depicted. Points are chosen in an order inversely proportional to the height of the distributions. Notice the penalty for large λ_1 for the Harris criterion. The fit distribution exhibits properties consistent with observed statistics including similar curvature (albeit of a simpler form) and asymptotic behavior along the λ_1 axis.

the distribution learned from the $\sigma = 2$ condition shown in figure 1 and this should not impact 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 the geometry of real physical entities. The following provides an example of the operation of the original Harris criterion and Shi criterion as compared with the likelihood based criterion as applied to some natural images with strong geometric structure as depicted in figure 3.

It appears that the likelihood based selection criterion results in more corners corresponding to the physical corners of objects than the Harris corner and additionally, better localization in some instances. Notice for example in the top figure that corners of the hexahedron are correctly selected by the likelihood based measure but are missed by the Harris criterion. The corner located at the back of the hexahedron marked with an asterisk (*) is symptomatic of one issue with the standard Harris criterion: This location consists of one eigenvalue that is very large combined with a smaller eigenvalue of intermediate size with a difference of two orders of magnitude. The κ value required to have this point selected would need to be < 0.01 which is well below the *recommended range* for the Harris threshold. This same phenomenon also appears to be behind the mislocalization of corners produced by the Harris measure as sufficiently sharp 2D image features may produce a λ_1 value that prohibits selection of the corner itself, but is sufficiently low for a location that is *near* the actual corner.

The images also reveal a significant shortcoming of the Harris and Shi proposals, that being, that both are susceptible to selecting regions for which no particularly strong features, or correspondence with image geometry exists. This is due in the case of the Harris detector, to its aversion to strong edge content, and in the case of the Shi operator, its ambivalence to λ_1 resulting in the acceptance of very flat regions. The boundary based on the statistical fit appears to implement a sensible tradeoff between the behavior of the Harris and Shi proposals. It is also interesting to note the similar behavior of the fit to the statistics and the Shi proposal, in particular for the middle row. This is perhaps unsurprising as the steep dropoff with increasing λ_2 implies that for a sufficiently large λ_1, the Shi proposal is a very good approximation of $p(\lambda_1, \lambda_2)$ and the least squares fit also reflects this consideration. The third pair of images strongly emphasizes the differences in behavior for the Shi and fit boundaries. While lower scoring points tend to reside on edges in the case of the fit distribution, they appear more akin to a cloud of noise on the rightmost poster for the Shi operator. The Harris operator (not shown) exhibits similar behavior and with a cloud of points that is distinct from the Shi operator. Although it is certain that much more rich descriptions of the $p(\lambda_1, \lambda_2)$ distribution might be established, it appears that the simplistic form put forth is sufficient to enforce the behavior corner $>$ edge $>$ flat/noise desired for modeling geometry and also arguably for matching purposes.

5. Quantitative Evaluation

On the basis of the discussion appearing in section 2 it is sensible to provide a demonstration that the proposal of likelihood based interest point selection yields desirable performance in terms of repeatability in a quantitative sense. There are a variety of 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 evaluates repeatability scores across a wide baseline for interest points chosen on the basis of the Harris, Shi and statistical corner criteria based on the Cambridge database [10].

5.1. Evaluation Protocol

There exist a few commonly used databases for performance evaluation [6, 10]. The former of these deals with the selection of affine regions and as such, the inclusion of a procedure for affine shape adaptation which results in a

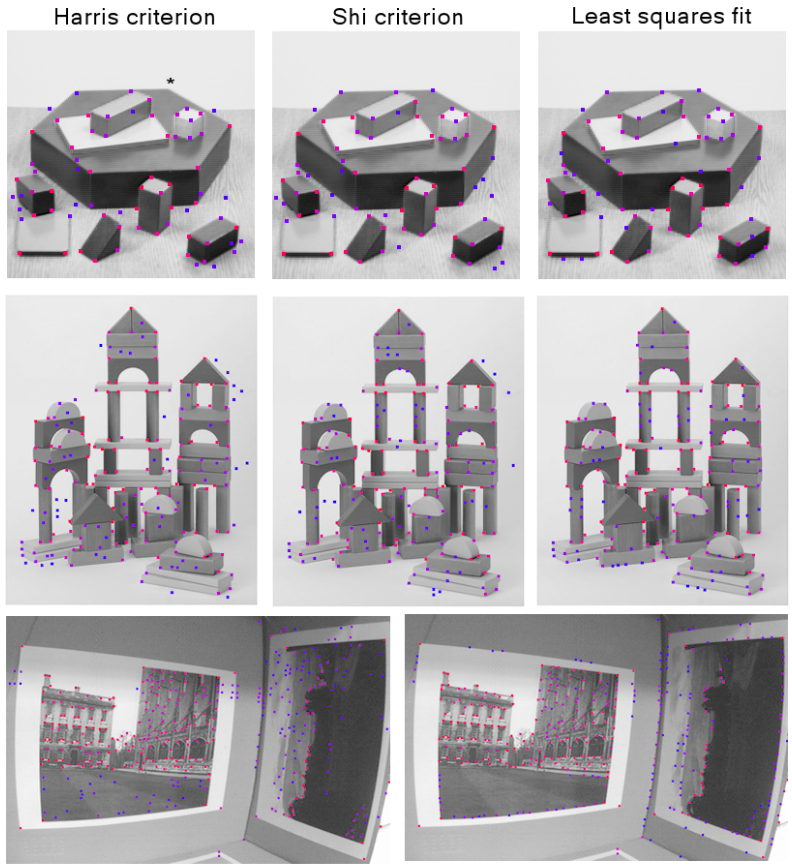

Figure 3. An example of corners selected by the original Harris criterion ($\kappa = 0.04$), the Shi-Tomasi criterion, and for the likelihood based criterion for a *classic* test image (top), for a similar example (middle). An image from the data set upon which quantitative evaluation was performed is presented on the bottom with the Shi output appearing on the left, and the fit output on the right. Points are color coded such that the strongest points are red, with a gradual shift to blue based on distance from the decision boundary.

confounded picture of the underlying selection boundaries. This is a consideration that is apparent in [1], for which very high dimensional polynomial fits yielded near perfect correspondence for the statistics within a specific range, but unpredictable behavior outside of this range. The non-convergence of points according to the affine shape adaptation procedure obscures the detriment to performance that arises out of such instability. For this reason, we have employed the evaluation protocol of Schmid et al. [11] on the data set of Rosten et al. [10] which allows direct comparison of interest points across a multitude of views. The determination of repeatability in this case is based on observing whether there exists a point x_i in image i that is within an ε neighborhood of a point x_j in image j subject to the homography H_{ij}. This was performed for 3 different sets of images and every pairwise comparison of selected points is considered.

A sample of the image set is depicted in figure 4 and consists of images from three distinct environments with views taken across a wide baseline. The image sets in figure 4 are referred to as the box (top row), maze (middle row) and junk (bottom row) data from hereon. In total, there are 14 images in the box data set, 15 in the maze data set and 8 in the junk data set. Performance evaluation is performed for pairwise correspondences across all possible pairs of images.

5.2. Results

Figure 5 demonstrates the repeatability scores for correspondences for the three sets of images for all possible pairwise correspondences. Performance results overall tend to favor the likelihood based criterion and Shi based criterion over the Harris proposal. Additionally, the likelihood based criterion performs better for the first several hundred points. This is suggestive of the fact that this criterion is the most stable in its decision boundary as the order in which points are selected is not reflected in the performance measure and will have the strongest influence on overall performance over the interval for which the number of points selected is relatively low. The frame in the bottom right

Figure 4. Test images from the study of Rosten et al. [10] the images are subjected to large changes of viewpoint. Examples of views from the three distinct sets are given.

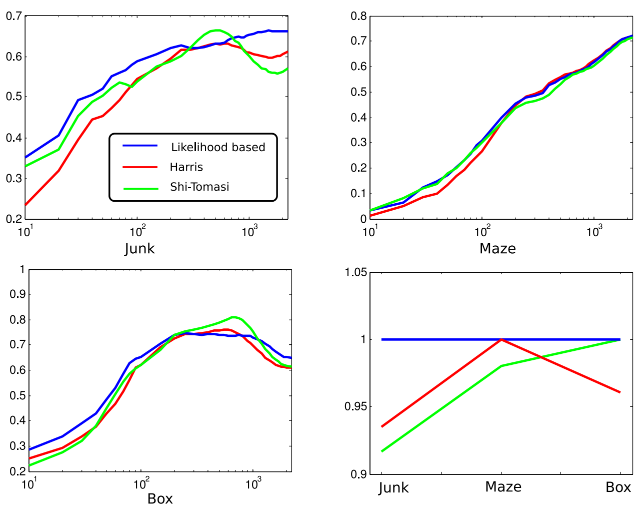

Figure 5. Repeatability scores for increasingly large numbers of chosen points for the junk (top left), maze (top right), and box (bottom left) data sets. The Bottom right frame depicts the ratio of the area under the curve for the three methods normalized by the area produced by the fit to the empirical data.

shows the area under the curve of the three criteria, normalized to the statistical criterion. It is apparent that the statistical criterion is equal or better to the other two proposals across all three sets of images subject to this criterion albeit, this is of course sensitive to the range of points considered, but it is encouraging that within a typical range, the statistical fit performs well. It is also worth noting that many matching applications typically employ on the order of 1000 points [5] and the statistical proposal is equal or above the other proposals within this range. The difference is also arguably greatest for the Junk class of scenes which notably has less strong features than the other two classes. This holds promise for performance for more naturalistic images in which such a large number of strong features is less likely to be present making sensitivity to noise more important. As a whole, the results are suggestive of the fact that a proposal fit to natural image statistics is a sensible strategy affording reasonable stability across the entire range of points and also apparently exhibiting greater correspondence with the geometry of the scene.

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 criteria, 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 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 than the simplistic form given. 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 and serves as a guide for the design of decision boundaries for interest operators in general. 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:

- The analysis presented calls into question the constrained form of the original Harris criterion since its design as a choice based on computational parsimony has lost much relevance. The analysis presented in this paper gives a sense of *why* the existing proposals work well, and also demonstrates *why* they fail. In particular for both the Harris and Shi proposal, there appears to be sensitivity to noise due to the aversion or ambivalence to λ_1.
- The minimum of the two eigenvalues of the autocorrelation matrix appears to provide a good approximation to the likelihood of the eigenvalue pair especially for large λ_1. An additional implication of this con-

sideration is that subject to the motivating critera, the Shi proposal is near optimal with respect to the desired distribution for sufficiently large λ_1. This is a condition that will almost always be met in the domain of tracking given the smaller number of feature points involved. This consideration affords a sense of why the Shi proposal presents *Good Features to Track* motivated by ecological statistics, beyond providing an autocorrelation matrix that is well conditioned. It may also be said that this affords a sense of why the Harris operator succeeds as it exhibits a boundary nearly asymptotic to the λ_1 axis up to some limit upon which noisy regions may be accepted in favor of strong edges. This is the reason for its strength, but also its weakness and as such argues strongly against the constrained form currently in common use.

- The proposal of Shi and Tomasi is heavily cited in the tracking literature, but is not considered in most of the performance evaluation papers despite being cited (e.g. [6]). The Shi-Tomasi definition appears to be underrepresented outside of tracking efforts and the current evidence calls for a systematic evaluation of the efficacy of this criterion with respect to matching performance and in multi-scale affine region selection given its apparent efficacy in producing repeatable points.
- The likelihood based criterion selects structure that is more consistent with the geometry of real objects making it amenable to the representation of image structure and the apparent failure of the Harris measure seems to be a result of penalizing locations for which two stronger eigenvalues are present, one of exceptional magnitude. This as mentioned, is also the apparent cause of its failure in mislocalizing strong corners. The Shi operator fails when a large number of points is required and relatively flat or noisy regions are chosen above edges due to its ambivalence to λ_1. A form based on the observed statistics appears to more strictly enforce the ordering $corner > edge > flat$ leading to quantitative stability for larger numbers of interest points, and also greater correspondence in the lower range suggesting that the ordering of selected points is more consistent.
- We have used a very simple choice of form to fit the eigenvalue PDF. It is likely that additional gains may be had in employing a form that is better able to capture the richness of the statistics in question however the extreme density at the origin makes fitting difficult in the absence of presupposing a rich form for the distribution consistent with its properties. That said, the present work is sufficient to point out some potential avenues for future improvements in selecting interest points, and also makes evident the tradeoff between noise and features of interest.
- An additional avenue for future consideration with respect to probabilistic determination of decision boundaries concerns 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. In this case one might construct a decision boundary that is optimal with respect to the joint consideration of $p(\lambda_1, \lambda_2)$ and p$(\Delta\lambda_1, \Delta\lambda_2)$.

As a whole, we have presented much food for thought in the domain of design of decision boundaries for interest points. A selection of Harris corners, even in simple form, motivated by structure observed in natural image statistics reveals some interesting aspects pertaining to the behavior of various decision criteria and also demonstrates where they fall short. The analysis presented raises many interesting questions pertaining to the selection of Harris corners, and the nature of interest operators in general and in addition, presents many fruitful directions for further research.

References

[1] N. D. Bruce and P. Kornprobst. Harris corners in the real world: A principled selection criterion based on ecological statistics. Technical Report RR-6745, INRIA, Dec 2008.

[2] 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.

[3] C. Harris and M. Stephens. A combined corner and edge detection. In *Proceedings of The Fourth Alvey Vision Conference*, pages 147–151, 1988.

[4] 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.

[5] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60:91–110, 2004.

[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 Sllskapet fr Bildanalys, 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] E. Rosten and T. Drummond. Machine learning for high-speed corner detection. In *European Conference on Computer Vision*, volume 1, pages 430–443, May 2006.

[11] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172, 2000.

[12] J. Shi and C. Tomasi. Good features to track. In *IEEE CVPR*, pages 593–600, 1994.