An Axiomatic Approach to Corner Detection

C. S. Kenney M. Zuliani B. S. Manjunath

Vision Research Lab
Department of Electrical and Computer Engineering
University of California, Santa Barbara

Abstract

This paper presents an axiomatic approach to corner detection. In the first part of the paper we review five currently used corner detection methods (Harris-Stephens, Förstner, Shi-Tomasi, Rohr, and Kenney et al.) for graylevel images. This is followed by a discussion of extending these corner detectors to images with different pixel dimensions such as signals (pixel dimension one) and tomographic medical images (pixel dimension three) as well as different intensity dimensions such as color or LADAR images (intensity dimension three). These extensions are motivated by analyzing a particular example of optical flow in pixel and intensity space with arbitrary dimensions.

Placing corner detection in a general setting enables us to state four axioms that any corner detector might reasonably be required to satisfy. Our main result is that only the Shi-Tomasi (and equivalently the Kenney et al. 2-norm detector) satisfy all four of the axioms

1. Introduction

Corner detection in images is important for a variety of image processing tasks including tracking, image registration, change detection, determination of camera pose and position and a host of other applications. In the following, the term "corner" is used in a generic sense to indicate any image feature that is useful for of establishing point correspondence between images.

Detecting corners has long been an area of interest to researchers in image processing. Some of the most widely used corner detection approaches rely on the properties of the averaged outer product of the image gradients:

$$L(\mathbf{x}, \sigma, g) = (G_\sigma * g)(\mathbf{x}) \quad (1a)$$

$$\mu(\mathbf{x}, \sigma, g) = \left(w * \nabla_{\mathbf{x}} L(\cdot, \sigma, g) \nabla_{\mathbf{x}}^T L(\cdot, \sigma, g)\right)(\mathbf{x}) \quad (1b)$$

where $L(\mathbf{x}, \sigma, g)$ is the smoothed version of the image g at the scale σ, and $\mu(\mathbf{x}, \sigma, g)$ is a 2×2 symmetric and positive semi-definite matrix representing the averaged outer product of the image gradients. The function w weights the pixels about the point $\mathbf{x}$. Förstner [1], in 1986 introduced a rotation invariant corner detector based on the ratio between the determinant and the trace of μ; in 1989, Noble [7] considered a similar measure in her PhD thesis. Rohr in 1987 [8] proposed a rotation invariant corner detector based solely on the determinant of μ. Harris and Stephens in 1988 [2] introduced a function designed to detect both corners and edges based on a linear combination of the determinant and the squared trace of μ, revisiting a previous work of Moravec [6]. This was followed by the corner detector proposed by Tomasi and Kanade in 1992 [11], and refined in 1994 in the well-known feature measure of Shi and Tomasi [10], based on the smallest eigenvalue of μ. All these measures create a value at each point in the image with larger values indicating points that are better for establishing point correspondences between images (i.e., better corners). Corners are then identified either as local maxima for the detector values or as points with detector values above a given threshold. All of these detectors have been used rather successfully to find corners in images but have the drawback that they are based on heuristic considerations. Recently Kenney et al. in 2003 [3] avoided the use of heuristics by basing corner detection on the conditioning of points with respect to window matching under various transforms such as translation, rotation-scaling-translation (RST), and affine pixel maps. Along similar lines Triggs [12] proposed a generalized form of the multi-scale Förstner detector that selects points that are maximally stable with respect to a certain set of geometric and photometric transformations. This paper extends the ideas contained in [3] and [13], where the corner detector function was defined as the reciprocal of the condition value. The condition theory framework allows one to extend corner detection to vector images such as color and LADAR and to images with different pixel dimensions such as signals (1D) and tomographic images (3D), similar to Rohr [9]. We present a set of four axioms that one might reasonably require a corner detector to satisfy, exploring the restrictions that the axioms impose on the set of allowable corner detectors. These restrictions are illustrated in a comparative study of the corner measures of Harris-Stephens, Förstner, Shi-Tomasi, Rohr and Kenney et al. . This paper is structured as follows. Section 2 introduces the corner detection problem in the context of single channel images. All of the detectors mentioned above rely on local gradient information to evaluate their respective detector measures. To motivate this reliance on the local gradients we present a simple thought experiment in the context of optical flow estimation. This experiment is also useful in our later extension of corner detectors to different pixel and intensity dimensions and it has the nice feature that it mirrors the results produced by condition theory without the attendant analytic complexity. Section 3 contains an axiomatic system for generalized corner detectors followed by a thorough discussion of the motivations and implications of the axioms. Finally the conclusions will be presented is Section 4. Almost all the proofs of the theorems and lemmas in the paper have been omitted for space reasons: the interested reader can refer to [4].

2. Notation and Motivation

Let $g = g(\mathbf{x})$ be the gray level intensity of an image at the image point $\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix}^T$. Let Ω be a window about $\mathbf{x}$ and define the gradient matrix A over this window by:

$$A(\mathbf{x}) \stackrel{\text{def}}{=} \begin{bmatrix} g_x^1 & g_y^1 \\ \vdots & \vdots \\ g_x^N & g_y^N \end{bmatrix}$$

where subscripts indicate differentiation and superscripts refer to the point location in the window Ω. To simplify the notation we will omit the dependence of A on $\mathbf{x}$.

The 2×2 gradient normal matrix is given by:

$$A^T A \stackrel{\text{def}}{=} \begin{bmatrix} \sum_{i=1}^N g_x^i g_x^i & \sum_{i=1}^N g_x^i g_y^i \\ \sum_{i=1}^N g_x^i g_y^i & \sum_{i=1}^N g_y^i g_y^i \end{bmatrix}$$

where the summation is over the window Ω about the point of interest. The gradient normal matrix $A^T A$ is the basis of all the corner detectors mentioned above. Note that this matrix can be obtained from (1b) under the assumption that w is the unit weight over the window Ω.

Why should a corner detector just depend on $A^T A$? We can motivate the reliance on the gradient normal matrix by looking at a specific problem in optical flow.

2.1. Optical Flow

Let $g = g(\cdot, \cdot, t)$ be an image sequence and suppose that a point of interest has time dependent coordinates $x = x(t)$ and $y = y(t)$. The optical flow problem is to discover the time evolution of x and y. In the standard approach this is done by making the assumption of constant brightness: $g(x(t), y(t), t) = c$, where c is a constant with respect to t. If we expand this constraint and neglect higher order terms we obtain $g_x dx + g_y dy + g_t dt = 0$ where subscripts denote differentiation. The previous equation can be rewritten in matrix form as $\begin{bmatrix} g_x & g_y \end{bmatrix} \mathbf{v} = -g_t \, dt$ where $\mathbf{v} = \begin{bmatrix} dx & dy \end{bmatrix}^T$ is the optical flow vector. This is one equation for the two unknowns dx and dy. To overcome this difficulty the standard approach is to assume that dx and dy are constant in a window about $\mathbf{x}$. This leads to the overdetermined set of equations:

$$\begin{bmatrix} g_x^1 & g_y^1 \\ \vdots & \\ g_x^N & g_y^N \end{bmatrix} \mathbf{v} = - \begin{bmatrix} g_t^1 \\ \vdots \\ g_t^N \end{bmatrix}$$

where we adopt a time scale in which $dt = 1$ and the superscripts indicate position within the window. More compactly we may write this as $A\mathbf{v} = \boldsymbol{\eta}$ where $\boldsymbol{\eta} = -\begin{bmatrix} g_t^1 & \dots & g_t^N \end{bmatrix}^T$. The least squares solution to this set of equations is obtained by multiplying both sides by A^T to obtain a square system and then multiplying by $(A^T A)^{-1}$ to get $\mathbf{v}_{computed} = (A^T A)^{-1} A^T \boldsymbol{\eta}$. A major problem with this approach is that some points give better estimates of the true optical flow than others. For example, if the image intensities in the window about $\mathbf{x}$ are nearly constant (uniform illumination of a flat patch) then $A \approx 0$ and the least squares procedure gives bad results.

2.2. A Thought Experiment

We can assess which points are likely to give bad optical flow estimates by a simple ansatz: suppose that the scene is static so that the true optical flow is zero: $\mathbf{v}_{exact} = 0$. If the images of the scene vary only by additive noise then $\boldsymbol{\eta}$ (the difference between frames) is just noise. The error in the optical flow estimate is given by $\mathbf{e} \stackrel{\text{def}}{=} \mathbf{v}_{exact} - \mathbf{v}_{computed}$, and we may write $\|\mathbf{e}\| \leq \|(A^T A)^{-1} A^T\| \, \|\boldsymbol{\eta}\|$. Thus we see that the term $\|(A^T A)^{-1} A^T\|$ controls the error multiplication factor; that is the factor by which the input error (the noise $\boldsymbol{\eta}$) is multiplied to get the output error (the error in the optical flow estimate). Large values of $\|(A^T A)^{-1} A^T\|$ correspond to points in the image where we cannot estimate the optical flow accurately in the presence of noise at least for the static image case. If we use the 2-norm together with the result that for any matrix M we have $\|M\|_2^2 = \lambda_{max}(MM^T)$, where $\lambda_{max}(MM^T)$ is the largest eigenvalue of MM^T, then we see that $\|(A^T A)^{-1} A^T\|_2^2 = \lambda_{max}((A^T A)^{-1}) = \frac{1}{\lambda_{min}(A^T A)}$ (where $\lambda_{min}(A^T A)$ indicates the smallest eigenvalue of $A^T A$). We conclude that the error multiplication factor for the 2-norm in the optical estimate for the static noise case is equal to $\frac{1}{\lambda_{min}(A^T A)}$. This motivates the use of the gradient normal matrix in feature detection since the ability to accurately determine optical flow at a point is intimately related to its suitability for establishing point correspondence between images (i.e., whether it is a good corner, see also [3]).

2.3. Corner Detection for Different Pixel and Intensity Dimension

The need to locate good points for tracking occurs in other setting besides images with two pixel dimensions and one intensity dimension. For example we may want to consider good matching points in signals (pixel dimension one) or tomographic medical images (pixel dimension three) or color images (intensity dimension is three) or hyperspectral images (intensity dimension much greater than one). In order to set up a framework for discussing corner detection for images with arbitrary pixel and intensity dimensions let $\mathbf{x} \stackrel{\text{def}}{=} \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ denote the pixel coordinates and $\mathbf{g} \stackrel{\text{def}}{=} \begin{bmatrix} g_1 & \dots & g_m \end{bmatrix}^T$ the intensity vector for the image. We use the optical flow method described above to set up a corner detection paradigm. It is worth noting that the results we obtain by this method are the same that would be obtained by applying condition theory [3] but are much easier to derive. That said, let $\mathbf{x} = \mathbf{x}(t)$ be a point of interest in a time dependent image $\mathbf{g} = \mathbf{g}(\cdot, t)$. We assume that this point has constant brightness over time $\mathbf{g}(\mathbf{x}(t), t) = \mathbf{g}(\mathbf{x}(t) + d\mathbf{x}, t + dt) = c$. Differentiating with respect to time we find that:

$$J\,\mathbf{v} = -\mathbf{g}_t \tag{2}$$

where we once again assumed that $dt = 1$ and the Jacobian matrix $J \in \mathbb{R}^{m \times n}$ has entries $[J]_{i,j} = \partial g_i / \partial x_j$, and:

$$\mathbf{v} = \begin{bmatrix} \frac{dx_1}{dt} & \dots & \frac{dx_n}{dt} \end{bmatrix}^T \qquad \mathbf{g}_t = \begin{bmatrix} \frac{dg_1}{dt} & \dots & \frac{dg_m}{dt} \end{bmatrix}^T$$

As before let $A = J$. If $A^T A$ is invertible then the least squares solution to (2) is given by $\mathbf{v} = (A^T A)^{-1} A^T (-\mathbf{g}_t)$. To illustrate this consider the problem for a signal (pixel dimension $n = 1$, intensity dimension $m = 1$). In this case the Jacobian is just the usual gradient of the signal: $J = dg/dx$ and the matrix $A^T A$ is invertible if the gradient is nonzero. Compare this with the case of an image (pixel dimension is two, intensity dimension is one). In this case the Jacobian is again the gradient $Jg = \nabla g = \begin{bmatrix} \partial g/\partial x & \partial g/\partial y \end{bmatrix}$ and the matrix $A^T A = \nabla g^T \nabla g$ is the outer product of the gradient row vector. Consequently the 2×2 matrix $A^T A$ for a grayscale image is rank deficient (its rank is at most 1) and so it is not invertible. This singularity disappears in the case of a color image. For example if $\mathbf{g} = \begin{bmatrix} R & G & B \end{bmatrix}^T$ then the rows of the Jacobian are the gradients of the red, green and blue channels:

$$J = \begin{bmatrix} \frac{\partial R}{\partial x} & \frac{\partial R}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \\ \frac{\partial B}{\partial x} & \frac{\partial B}{\partial y} \end{bmatrix} = \begin{bmatrix} \nabla R \\ \nabla G \\ \nabla B \end{bmatrix}$$

In this case the 2×2 matrix $A^T A = \nabla R^T \nabla R + \nabla G^T \nabla G + \nabla B^T \nabla B$ is the sum of the outer products of the three color channel gradient row vectors. Consequently it is invertible if any two of the channels have independent gradient vectors. In general we find that $J^T J = \sum_{i=1}^{m} (\nabla g_i)^T \nabla g_i$. From this we conclude that the gradient normal matrix $J^T J$ is $n \times n$ where n is the pixel dimension and has rank at most m where m is the intensity dimension. It is not invertible if the pixel dimension exceeds the intensity dimension: $n > m$. If the pixel dimension is larger than the intensity dimension then we may overcome the non-invertibility of $A^T A$ by making the additional constraint that the optical flow is locally (i.e. in a window) constant. In this case the equation (2) holds over the window and the least squares solution is obtained by stacking these sets of equations into a large system:

$$A\mathbf{v} = \begin{bmatrix} J^1 \\ \vdots \\ J^N \end{bmatrix} \mathbf{v} = - \begin{bmatrix} d\mathbf{g}^1/dt \\ \vdots \\ d\mathbf{g}^N/dt \end{bmatrix} = \boldsymbol{\eta}$$

As seen earlier for the grayscale image case even in this general setting the term $\|(A^T A)^{-1} A^T\|$ controls the error multiplication factor [4]; this motivates the role of $A^T A$ in corner detector for the general problem of arbitrary pixel and intensity dimensions. For the purposes of interpretation it is helpful to rewrite $A^T A$ as:

$$A^T A = \sum_{j=1}^{N} \sum_{i=1}^{m} \left(\nabla g_i^j\right)^T \nabla g_i^j = \sum_{j=1}^{N} \left(J^j\right)^T J^j \tag{3}$$

That is, $A^T A$ is the sum over the window of the outer products of the gradient vectors of each intensity channel.

3. Axioms for Corner Detectors

In order to formulate the axioms that a reasonable corner detector might be required to satisfy we need the following definitions.

Definition 1 *A (local)* corner detector *for an image with pixel dimension n and intensity dimension m is a real-valued function f of $A^T A$ as given by* (3) *for the pixel location $\mathbf{x}$ and a given window Ω about $\mathbf{x}$.*

To compare detector values for different pixel and/or intensity dimensions (Axioms 1 and 2 below) we assume that the corner detector is defined for positive semi-definite matrices of differing sizes.

Definition 2 *$S_1 \leq S_2$ if $S_2 - S_1$ is positive semi-definite.*

Definition 3 *Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $A^T A$ at $\mathbf{x}$. We say that a set of points X in the image $\mathbf{g}$ has constant* eigen-energy *with respect to the q-norm if $\lambda_1^q + \dots + \lambda_n^q$ is constant over $\mathbf{x} \in X$.*

Definition 4 *A point* $\mathbf{x}$ *is* isotropic *(with respect to the image* $\mathbf{g}$*) if the eigenvalues of the gradient normal matrix are all equal:* $\lambda_1 = \lambda_2 = \ldots = \lambda_n$.

Before presenting the set of axioms we also need to introduce the concept of image restriction to an affine space in the pixel domain and image projection onto a subspace in the intensity domain. Consider a multidimensional and multichannel image $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$ such that a point in the affine space $\mathcal{P}$ of dimension d passing through the point $\mathbf{x}$ can be written as $P\mathbf{p}+\mathbf{x}$ for some suitable $\mathbf{p}$ (P is an $n \times d$ orthogonal matrix). The Jacobian of $\mathbf{g}$ is related to the Jacobian of its restriction to $\mathcal{P}$ according to $J_{\mathbf{p}} = J_{\mathbf{x}}\, P$ and therefore the normal matrix becomes $A^TA \to P^TA^TAP$. Similarly, if we consider a subspace $\mathcal{Q}$ in the intensity space that is spanned by the orthonormal columns of Q, the projection onto $\mathcal{Q}$ is given by $\mathbf{h} = Q^T\mathbf{g}$. In this case the Jacobians are related according to $J_Q = Q^TJ$ and consequently the normal matrix becomes $A^T\bar{Q}\bar{Q}^TA$ where:

$$\bar{Q} = \begin{bmatrix} Q & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & Q \end{bmatrix} \in \mathbb{R}^{Nm \times Nq}$$

Note that the block matrix $\bar{Q}$ still has a set of orthonormal columns. We have now all the tools that are needed to present the following set of axioms.

Axiom 1 *Let* $P \in \mathbb{R}^{n\times d}$ *with* $d \le n$ *be a matrix with orthonormal columns. Then* $f(A^TA) \le f(P^TA^TAP)$ *and equality is achieved if and only if* $\mathbf{x}$ *is a point of isotropy (*isotropy condition*) or* $d = n$ *(*rotation invariance condition*).*

Axiom 2 *Let* $\bar{Q} \in \mathbb{R}^{Nm\times Nq}$ *with* $q \le m$ *be a matrix with orthonormal columns. Then* $f(A^TA) \ge f(A^T\bar{Q}\bar{Q}^TA)$.

Axiom 3 *If* $A_1^TA_1 \le A_2^TA_2$ *then* $f(A_1^TA_1) \le f(A_2^TA_2)$.

Axiom 4 *The corner detector over a set of constant eigenenergy points attains its maximum value at a point of isotropy.*

We now consider motivation and consequences for these axioms. For this purpose we introduce the following generalized detectors:

- Generalized Harris-Stephens corner detector:

$$f_{HS} \overset{\text{def}}{=} \det(A^TA) - \alpha\left(\text{trace}(A^TA)\right)^n = \prod_{i=1}^{n}\lambda_i - \alpha\left(\sum_{i=1}^{n}\lambda_i\right)^n \quad (4)$$

  where α is a user supplied constant that controls the sensitivity of the corner detector.

- Generalized Förstner corner detector:

$$f_F \overset{\text{def}}{=} \frac{1}{\text{trace}\left((A^TA)^{-1}\right)+\varepsilon} = \frac{1}{\varepsilon + \sum_{i=1}^{n}\frac{1}{\lambda_i}} \quad (5)$$

  We note in passing that nf_F is equal to the harmonic mean of the eigenvalues of the gradient normal matrix A^TA (provided $\varepsilon = 0$). Also if any of the λ_i's is zero then we set $f_F = 0$.

- Generalized Shi-Tomasi corner detector:

$$f_{ST} \overset{\text{def}}{=} \lambda_{min}(A^TA) \quad (6)$$

- Generalized modified Rohr[1] corner detector:

$$f_R \overset{\text{def}}{=} \sqrt[n]{\det(A^TA)} \quad (7)$$

- Generalized Kenney et al. corner detector for the Schatten p-norm:

$$f_{K,p} = \frac{1}{\left\|(A^TA)^{-1}\right\|_p} = \frac{1}{\left(\sum_{i=1}^{n}\frac{1}{\lambda_i{}^p}\right)^{\frac{1}{p}}} \quad (8)$$

  As a matter of notation we will refer to $f_{K,p}$ as the p-norm condition detector.

Henceforth we will assume that the eigenvalues of A^TA are arranged in non increasing order, i.e. $\lambda_1 \ge \ldots \ge \lambda_n$.

Lemma 1 *The Förstner, Shi-Tomasi, and Kenney et al. corner detectors are equivalent modulo the choice of a suitable matrix norm. Rohr's modified detector is equivalent to Kenney et al. detector in a limit sense (via a normalization constant).*

$$\begin{aligned} f_F &= f_{K,1} \quad \textit{(provided } \varepsilon = 0\textit{)} \\ f_{ST} &= f_{K,\infty} \\ f_R &= \lim_{p\to 0}\frac{1}{\sqrt[p]{2}} f_{K,p} \end{aligned}$$

where $f_{K,\infty} \overset{\text{def}}{=} \lim_{p\to\infty} f_{K,p}$.

3.1. Axiom 1

To motivate this axiom, consider an image which is black to the left of the center line and white to the right of the center line (see Figure 1). Such an image has an aperture effect in that we may be able to determine left-right motion but not up-down motion. That is any point $\mathbf{x}$ on the center line is not suitable as a feature for full motion detection.

[1] In [8], Rohr studied the detector obtained from determinant of A^TA: to simplify the equivalence result that will be stated in Lemma 1 we will instead consider the modified version $\sqrt{\det(A^TA)}$.

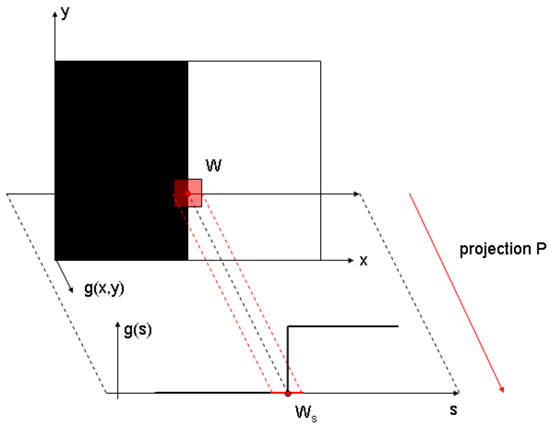

Figure 1. The figure shows the projection of the pixel space on a lower dimensional subspace $\mathcal{P}$ ($d = 2$).

This is seen in the eigenvalues of the gradient normal matrix $A^T A$:

$$\min(\lambda_1, \lambda_2) = 0 \tag{9}$$

Thus we get a zero value for the Förstner, Shi-Tomasi, modified Rohr and p-norm condition detectors; the Harris-Stephens detector gives a negative value for this example.

Now suppose that we pass a line through $\mathbf{x}$ and consider the signal of intensity values from the original image along this line. This signal is piecewise constant with a step as it crosses through $\begin{bmatrix} x & y \end{bmatrix}^T$. Thus it has a good feature for tracking at $\mathbf{x}$; the restriction to a lower dimensional subspace has improved the corner detection properties of the point. As a technical note if the subspace line that we choose through $\mathbf{x}$ is vertical then no step will appear and the point is still not suitable as a feature tracking point. However this does not violate the spirit of Axiom 1 since the point was already unsuitable as a corner in the original (higher dimensional) setting.

Lemma 2 *Any p-norm condition generalized corner detector*[2] *satisfies the condition $f(A^T A) \leq f(P^T A^T AP)$. The Shi-Tomasi corner detector also satisfies the isotropy condition, whereas the Förstner detector (and consequently the 1-norm condition detector) does not. The generalized Harris-Stephen and modified Rohr corner detector violate the condition $f(A^T A) \leq f(P^T A^T AP)$.*

Remark 1 *We have included the isotropic equality requirement in Axiom 1 in order to ensure that if the point $\mathbf{x}$ is a local maximum for the corner detector then it remains a local maximum if we restrict the detector to a subspace through $\mathbf{x}$. The reason for this is that we may want to attain efficiency of detection by for example using a 1D corner detector in say the x-direction; we could then cull the points which are poor 1D corners and then do a full corner detector evaluation at the remaining points in the image. If the detector satisfies Axiom 1 then we would be assured that local maxima for the full detector were not eliminated during the preliminary 1D sweep.*

When we choose d to be equal to n this axiom states the intuitive fact that a corner should remain a corner independent of orientation or reflection of the image. This fact can be expressed requiring that $f(A^T A) = f(P^T A^T AP)$ for any orthogonal matrix P. As an immediate consequence we have:

Lemma 3 *Any corner detector satisfying Axiom 1 for $d = n$ depends only on the eigenvalues $\lambda_1, \ldots, \lambda_n$ of $A^T A$. That is (with a slight abuse of notation) we may write $f(A^T A) = f(\lambda_1, \ldots, \lambda_n)$.*

Remark 2 *Lemma 3 is similar in spirit to Von Neumann's classic result on the equivalence of unitarily invariant norms and symmetric gauge functions (see [5]).*

3.2. Axiom 2

To motivate this axiom we consider the case of color images (say RGB) and compare with the restriction to one color channel (say R). If we work only at the point $\mathbf{x}$ (just take the window Ω to be the point $\mathbf{x}$) then:

$$\begin{aligned} A^T A &= \nabla R^T \nabla R + \nabla G^T \nabla G + \nabla B^T \nabla B \\ A^T Q Q^T A &= \nabla R^T \nabla R \end{aligned}$$

where $Q = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$. Clearly we have $A^T Q Q^T A \leq A^T A$ so that by Axiom 3 we want $f(A^T Q Q^T A) \leq f(A^T A)$ which is what Axiom 2 requires.

Lemma 4 *The generalized Förstner and Shi-Tomasi corner detector (and consequently the 1-norm and ∞-norm condition detector) and the Rohr corner detector satisfy Axiom 2. The generalized Harris-Stephen corner detector violates Axiom 2.*

3.3. Axiom 3

The matrix $A^T A$ provides a measure of both the strength of the intensity gradients and their independence. This can be encapsulated by the natural ordering on symmetric matrices. Thus the condition $A_1^T A_1 \leq A_2^T A_2$ in Axiom 3 means that the gradient vectors at $\mathbf{x}_2$ are stronger and/or more independent than those at $\mathbf{x}_1$ where $A_1 = A(\mathbf{x}_1)$ and $A_2 = A(\mathbf{x}_2)$.

[2]Note that Rohr's detector cannot be considered a p-norm condition generalized corner detector: in fact its value is zero if any of the eigenvalues is zero.

Lemma 5 *A corner detector satisfying the rotation invariance condition of Axiom 1 and Axiom 3 is nondecreasing in* $\lambda_1, \ldots, \lambda_n$.

Lemma 6 *The Förstner, Shi-Tomasi, modified Rohr and* p*-norm condition detectors are nondecreasing with respect to* $\lambda_1, \ldots, \lambda_n$*. However this is not true for the generalized Harris-Stephens detector.*

3.4. Axiom 4

If the matrix $A^T A$ has a large value of $\mathbf{v}^T A^T A \mathbf{v}$ for a vector $\mathbf{v}$ then it is well-conditioned for point matching with respect to translational shifts from $\mathbf{x}$ in the direction $\mathbf{v}$. As a directional vector $\mathbf{v}$ moves over the unit sphere the values of $\mathbf{v}^T A^T A \mathbf{v}$ pass through all the eigenvalues $\lambda_1, \ldots, \lambda_n$ of $A^T A$. This means that if one eigenvalue is smaller than the rest then the corresponding eigenvector $\mathbf{v}$ is a direction in which the corner is less robust (in the sense of point matching conditioning) than in the other eigenvector directions. From this we see that Axiom 4 can be interpreted as the requirement that the best corner (as measured by the corner detector function f) subject to the restriction of constant eigen-energy $\lambda_1^q + \cdots + \lambda_n^q = c$ for some $q > 1$ is that corner that doesn't have a weak direction: all the unit norm directional vectors $\mathbf{v}$ yield the same value for $\mathbf{v}^T A^T A \mathbf{v}$. That is we must have $\lambda_1 = \ldots = \lambda_n$. This reasoning motivated the Definition 4 of an isotropic point in the image. In order to test the generalized corner detectors for compliance with Axiom 4 it is helpful to rewrite the axiom as follows: over the set of eigenvalues of constant energy $\lambda_1^q + \cdots + \lambda_n^q = c$ for a given $q \geq 1$ and a constant c, the maximum of the corner detector is attained at $\lambda_1 = \lambda_2 = \cdots = \lambda_n = c/n^{1/q}$. (Note that we have restated Axiom 4 in this way to avoid complications resulting from images in which the set of points for a given eigen-energy may not contain all possible combinations of eigenvalues at that energy.)

Lemma 7 *The generalized Förstner, Shi-Tomasi, modified Rohr and* p*-norm condition detectors satisfy Axiom 4.*

Remark 3 *We can illustrate Axiom 4 and the above lemma by taking* $q = 1$ *in the eigen-energy measure. In this case we have* $\lambda_1 + \lambda_2 + \cdots + \lambda_n = c$*. That is the trace of* $A^T A$ *is constant. Moreover, using the linearity of the trace operator together with the property that: trace*$(M_1 M_2) =$ *trace*$(M_2 M_1)$ *for any compatibly dimensioned matrices* M_1 *and* M_2*, we find that trace*$(A^T A) = \sum_{j=1}^{N} \sum_{i=1}^{m} \|\nabla g_i^j\|^2$*. We note that the last term is just the sum of the squares of the norms of the intensity gradients over the window* Ω *about* $\mathbf{x}$*. This means that the condition that the eigen-energy is equal to* c *for* $q = 1$ *is the same as requiring the average of the squares of the gradient norms to also be constant. For example consider the case of a black square in the lower left of an otherwise white image (see Figure 2-a). Let us look at the constant eigen-energy line for the trace norm (*$q = 1$*) starting at the lower right boundary of the black square (see Figure 2-b). At this point we have that the two eigenvalues of* $A^T A$ *are* $\lambda_1 = c$*,* $\lambda_2 = 0$*. This remains fixed as we move upward along the line of constant eigen-energy. As we near the corner the line of constant eigen-energy curves inward and we reach a point where* $\lambda_1 = \lambda_2 = c/2$*. Continuing on the constant energy curve to the left we return to the state where the larger eigenvalue is equal to* c *and the smaller is equal to 0. Axiom 4 in this example requires the corner detector along this constant energy curve to be maximized at the point where* $\lambda_1 = \lambda_2 = c/2$*. This is also the point of closest approach of the curve to the true corner.*

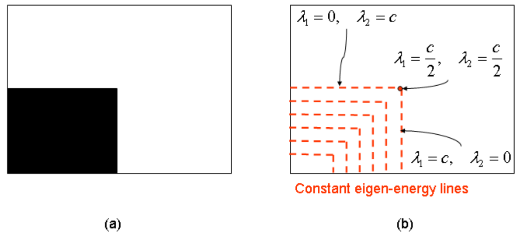

Figure 2. The left figure shows the test image whereas the right image shows the corresponding eigen-energy lines.

3.5. Shi-Tomasi Detectors

It has been shown that the Shi-Tomasi corner detector satisfies all the proposed axioms. This fact can be generalized using the following definition and lemma.

Definition 5 *A corner detector that is a function of* $\lambda_{min} = \min_{1 \leq i \leq n} \lambda_i$*, where* $\lambda_1, \ldots, \lambda_n$ *are the eigenvalues of* $A^T A$*, is called a* Shi-Tomasi type *detector.*

Lemma 8 *Let* f *be a corner detector in the sense of Definition 1. If* f *satisfies Axiom 1 and Axiom 3 then* f *is a Shi-Tomasi type detector.*

4. Conclusions

In this paper we have presented an axiomatic approach to corner detection. Our original purpose was to compare currently used corner detectors including the Harris-Stephens, Förstner, Shi-Tomasi, Rohr, and the p-norm condition detectors. However we found that by extending the definition of these detectors to include image spaces of differing pixel and intensity dimensions we were able to set up a general framework of four axioms that such detectors should satisfy. Motivation has been provided for each of these axioms

Table 1. Compliance of the generalized corner detectors with the proposed axioms.

	Axiom 1			Axiom 2	Axiom 3	Axiom 4
	$f(A^T A) \leq f(P^T A^T AP)$	Rotation Invariance	Isotropy Condition			
Harris-Stephens	×	✓	×	×	×	only for $n = 2$
Förstner (1-norm condition)	✓	✓	×	✓	✓	✓
Shi-Tomasi (∞-norm condition)	✓	✓	✓	✓	✓	✓
Modified Rohr	×	✓	×	✓	✓	✓

and they may serve as a basis either individually or collectively for testing future detection schemes. In the process of our analysis we also demonstrated that the Shi-Tomasi detector was equivalent to the ∞-norm condition detector and that the Förstner detector was equivalent to the 1-norm condition detector.

In our comparison of the five current detectors we showed that only the Shi-Tomasi (and equivalently the ∞-norm condition detector) was compliant with all four axioms. In contrast, the Harris-Stephens detector failed to satisfy Axioms 1 (except for the rotation invariance condition), 2 and 3, the Förstner detector failed to satisfy the isotropy condition in Axiom 1 and the modified Rohr detector failed to satisfy the basic condition and the isotropy condition of Axiom 1 . These considerations are summarized in Table 1.

Acknowledgments. The authors would like to thank prof. Shiv Chandrasekaran, prof. Karl Rohr, prof. Bill Triggs, and Dmitry Fedorov for the helpful discussions and suggestions. This project was supported by the grants ONR #N00014-04-1-0121 and ONR #N00014-02-1-0318.

References

[1] W. Förstner. A feature based correspondence algorithm for image matching. In *International Archives of Photogrammetry and Remote Sensing*, volume 26, pages 150–166, 1986.

[2] C. Harris and M. Stephens. A combined corner and edge detector. In M. M. Matthews, editor, *Proc. of the 4th ALVEY vision conference*, pages 147–151, University of Manchester, England, Septemeber 1988.

[3] C. Kenney, B. Manjunath, M. Zuliani, G. Hewer, and A. Van Nevel. A condition number for point matching with application to registration and post-registration error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(11):1437–1454, November 2003.

[4] C. S. Kenney, M. Zuliani, B. S. Manjunath, L. Peterson, and D. Schwartzwald. An axiomatic approach to corner detection. Technical Report TR-VRL-11-2004, Vision Research Lab, Department of Electrical and Computer Engineering, University of California, Santa Barbara, 2004.

[5] A. Lewis. The mathematics of eigenvalue optimization. *Mathematical Programming*, 97:155–176, 2003.

[6] H. Moravec. Obstacle avoidance and navigation in the real world by a seeing robot rover. Technical Report CMU-RI-TR-3, Carnegie-Mellon University, Robotics Institute, September 1980.

[7] A. Noble. *Descriptions of Image Surfaces*. PhD thesis, Department of Engineering Science, Oxford University, 1989.

[8] K. Rohr. Localization properties of direct corner detectors. *Journal of Mathematical Imaging and Vision*, 4(2):139–150, 1994.

[9] K. Rohr. On 3d differential operators for detecting point landmarks. *Image and Vision Computing*, 15(3):219–233, 1997.

[10] J. Shi and C. Tomasi. Good features to track. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pages 593–600, Seattle, Washington, June 1994.

[11] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography—a factorization method. *International Journal on Computer Vision*, 9(2):137–154, November 1992.

[12] B. Triggs. Detecting keypoints with stable position, orientation, and scale under illumination changes. In *Proc. of the 8th European Conference on Computer Vision*, volume 4, pages 100–113, 2004.

[13] M. Zuliani, C. Kenney, and B. Manjunath. A mathematical comparison of point detectors. In *Proc. of the 2nd IEEE Workshop on Image and Video Registration*, 2004.