Singularity Detection in Images Using Dual Local Autocovariance

Wojciech Czaja[1]
Department of Mathematics, University of Maryland, College Park, Maryland
E-mail: wojtek@math.umd.edu

and

Mladen Victor Wickerhauser[2]
Department of Mathematics, Washington University, St. Louis, Missouri
E-mail: victor@math.wustl.edu

Differences in the eigenvalues of an autocovariance matrix indicate directions at which the local Fourier power spectrum of a function is slowly decreasing. This provides a technique to discriminate edge-like singularities from other features in images.

Key Words: local Fourier power spectrum, edge detection, moments, Morrey-Campanato spaces, nondifferentiable regularity

1. INTRODUCTION

In [1, 8], we used the autocovariance matrix as a substitute for the Jacobian to investigate geometric properties of functions, such as the local dimension of the range. Specializing to images, this idea may be used, for instance, to detect points, edges, and other features. In this article, we concentrate on edges with the relatively weak regularity defined by Morrey and Campanato [4].

The literature on the subject of edge detection is enormous. We mention just a few such algorithms which are related to our own: Aron and Kurz [5], linear hypothesis testing of variances in small windows to detect lines and edges; Duda and Hart [2], Hough transform for detection of lines and curves in images; Mallat and Zhong [6], edge detection from wavelet maxima. Many other ways to use Fourier

[1]Permanent address: Instytut Matematyczny, Uniwersytet Wrocławski, Pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland

[2]Supported in part by NSF grant DMS-0072234

transforms to detect singularities are described in [3]. These algorithms separate into a *local* part, such as discrete Laplacian or Sobel difference filtering, followed by a *global* part such as template matching that recognizes edge-like groups of pixels. The local operation is based on approximate differentiation, either by finite differences as in the Sobel detector, or after transformation as in the Fourier and wavelet methods. It assigns a large value at singular points of the image, and a small value at smooth points. It typically produces too many candidate edge points, which must then be screened for membership along some line or curve. Candidate edge points are especially overabundant in noisy images.

To improve this situation, we introduce a local operation that produces a large value at a point only if it and a few lined-up neighbors are singular points. Both single-point singularites and nonsingular points of the image will produce small values. This will reduce the number of candidates to be checked by the global follow-up, especially in moisy images. Our local step has complexity comparable to filtering or transformation, but the global part's complexity grows with the number of candidates, so we expect our method to require less total computational effort.

We begin with a classical observation: whenever a function is not smooth at some point, then the power in its Fourier transform localized near that point will be slowly decreasing at high frequencies. But then, if the singularity has a direction, such as the normal direction to an edge discontinuity, the decrease will be particularly slow in that direction. This slow decrease creates a large variance in the slow direction, if we treat the local Fourier power spectrum as a probability density. By contrast, the variance in the other directions, in which the Fourier transform decreases rapidly, will be smaller. These variances are the two eigenvalues of the 2×2 autocovariance matrix, or equivalently the second-moment matrix, of the localized Fourier power spectrum.

Our new technique is to recognize the edge-like nonsmooth points of a function by the differences between these two eigenvalues. For theoretical analysis, we compute the limit eigenvalues as the localization shrinks to the point of interest. The "edginess" of a point will be a function of the ratio or difference of these limit eigenvalues, with bigger differences or ratios giving more "edginess." We will further show that the eigenvector of the larger eigenvalue will be normal to the edge, when such a normal exists.

The eigenvalues might be the same because they are small and equal, or large and equal. The first case arises at a point of smoothness, the second at a point singularity. Our technique assigns low "edginess" in both cases, and therefore differs from the differentiation-based edge detectors. Drawing a conclusion from two eigenvalues specializes our earlier work, in which we estimated the local rank of a complicated function from the number of relatively large eigenvalues of the autocovariance matrix.

We have implemented our algorithms in Standard C, and the source codes are freely available from the ACHA Software Distribution Web Site. The edginess functions of five example images were computed with this software and are displayed at the end of this article. Readers are invited to experiment with parameter variations using our five images, which are also on the web site. The codes may be used without modification on any other images in the simple PGM format.

2. MOTIVATION, THEORY AND EXAMPLES

Our motivation comes from probability theory. Suppose $\phi : \mathbf{T} \to \mathbf{C}$ is a function on the unit circle $\mathbf{T}$. We define a map $\Phi : \mathbf{T} \to \mathbf{C}^2$ as follows:

$$\Phi(\theta) = \phi(\theta)(\cos\theta, \sin\theta),$$

for $\theta \in \mathbf{T}$. Regarding Φ as a random vector on the probability space T, and ignoring normalization, we may define the 2×2 *autocovariance matrix*:

$$E(\Phi^*\Phi)_{ij} = \int_{\mathbf{T}} \Phi_i^* \Phi_j = \int_0^{2\pi} |\phi(\theta)|^2 \tau_i(\theta)\tau_j(\theta)\, d\theta,$$

where $i, j \in \{1, 2\}$, $\tau_1(\theta) = \sin\theta$, and $\tau_2(\theta) = \cos\theta$. This is a symmetric matrix. If we apply this matrix to a vector v, then $\langle E(\Phi^*\Phi)v, v\rangle = \int_{\mathbf{T}} |\langle \Phi, v\rangle|^2$, so in particular the supremums over all v with $\|v\| = 1$ are equal. A vector realizes this supremum if and only if it is a unit eigenvector of the largest eigenvalue of $E(\Phi^*\Phi)$. Such a vector always exists, and it will be our approximation to the point at which $|\phi| = \|\Phi\|$ attains its maximum. For example, in the special case that ϕ is highly concentrated near the point $\theta_0 \in \mathbf{T}$, we see that $E(\Phi^*\Phi)$ will be approximately proportional to

$$\begin{pmatrix} \cos^2\theta_0 & \cos\theta_0 \sin\theta_0 \\ \cos\theta_0 \sin\theta_0 & \sin^2\theta_0 \end{pmatrix},$$

for which the vector that points from the origin to θ_0 is an eigenvector of the largest eigenvalue.

Instead of the probability space $\mathbf{T}$, we may integrate over a disc B to get a version of the autocovariance matrix studied in [8]:

$$E(\Phi^*\Phi)_{ij} = \int_B |\phi(\xi)|^2 \xi_i \xi_j \, d\xi,$$

for $i, j \in \{1, 2\}$, and $\phi \in L^2(B)$.

To get localized information on the singularities of a function $f : \mathbf{R}^2 \to \mathbf{R}$, we will take ϕ to be the Fourier transform of f, after multiplication by a smooth cutoff function, or "bump," concentrated around the point of interest. This bump should be radial, to avoid introducing directional bias.

Definition 2.1. Fix a nonzero radial function $g : \mathbf{R}^2 \to \mathbf{R}$ in the Schwartz class, centered at the origin, and fix $\epsilon > 0$. Then for each polynomially bounded function $f : \mathbf{R}^2 \to \mathbf{R}$, and each point $x^0 \in \mathbf{R}^2$, define the *dual local autocovariance matrix of f at x^0* to be the 2×2 matrix

$$E_{\epsilon,g}(f; x^0)_{ij} = \int_{B(0,1/\epsilon)} \xi_i \xi_j \left| \widehat{g_\epsilon f}(\xi) \right|^2 d\xi, \qquad i, j \in \{1, 2\},$$

where $g_\epsilon(\xi) = g\left(\dfrac{\xi - x^0}{\epsilon}\right)$.

By these assumptions, gf is integrable, so $\widehat{gf}$ is bounded and continuous and the matrix coefficients are well defined. It is the real, symmetric second moment matrix of the unnormalized probability density function $|\widehat{g_\epsilon f}|^2$. If f is nonzero in a neighborhood of x^0, then the matrix will be positive definite. Our technique is to use differences and ratios of its eigenvalues to define the "edginess" of the function f, at the point x^0.

The dual local autocovariance matrix can also be defined for certain singular measures and distributions. For example, let f be the Dirac delta measure supported at y^0, and let g_ϵ be centered at x^0 as in Definition 2.1. Then $|\widehat{g_\epsilon f}(\xi)|^2 = |g_\epsilon(y^0)|^2$, which tends to $0+$ as $\epsilon \to 0$ if $x^0 \neq y^0$, but remains constantly 1 as $\epsilon \to 0$ if $x^0 = y^0$. In either case, $E_{\epsilon,g}(f; x^0)$ tends to a multiple of the identity as $\epsilon \to 0$, so both the ratio and the normalized difference of the eigenvalues is everywhere the same in the limit. Hence, either definition of edginess ignores point singularities.

2.1. Straight edges

After a change of variables, it is easy to see that:

$$E_{\epsilon,g}(f; x^0)_{ij} = \int_B \xi_i \xi_j \left| \int_{\mathbf{R}^2} g\left(x + \frac{\epsilon - 1}{\epsilon} x^0\right) f(\epsilon x) \exp(-2\pi i x \cdot \xi)\, dx \right|^2 d\xi. \quad (1)$$

We can therefore compute the dual local autocovariance matrix explicitly for the homogeneous example $f = \mathbf{1}_L$, the characteristic function of the left half-plane $L = \{(x_1, x_2) : x_1 \leq 0\}$. The graph of f presents an edge along the line $x_1 = 0$. We choose the radial function $g(x) = \exp(-\pi|x|^2)$, for ease of computing its Fourier transform.

Lemma 2.1. *Matrix $E_{\epsilon,g}(\mathbf{1}_L; 0)$ has distinct eigenvalues λ_1, λ_2 that correspond to the eigenvectors $(1,0), (0,1)$ and satisfy $\lambda_1 > \lambda_2 > 0$, independently of ϵ.*

Proof. Since $\mathbf{1}_L$ is homogeneous of degree 0, and $x^0 = 0$, we can eliminate ϵ in Equation 1:

$$E_{\epsilon,g}(\mathbf{1}_L; 0)_{ij} = \int_B \xi_i \xi_j |\widehat{g\mathbf{1}_L}(\xi)|^2 \, d\xi.$$

Hence the eigenvalues λ_1, λ_2 do not depend on ϵ.

Now $E_{\epsilon,g}(\mathbf{1}_L; 0)_{12} = E_{\epsilon,g}(\mathbf{1}_L; 0)_{21} = 0$ by the symmetry of $\widehat{g\mathbf{1}_L}$ with respect to ξ_2, so the matrix is diagonal. The two eigenvalues, corresponding to eigenvectors $(1,0), (0,1)$, are therefore $\lambda_i = E_{\epsilon,g}(\mathbf{1}_L; 0)_{ii} = \int_B \xi_i^2 |\widehat{g\mathbf{1}_L}(\xi)|^2 \, d\xi$, for $i = 1, 2$. Since g is nonzero, it is evident that λ_1 and λ_2 are positive. On the other hand, $|\widehat{g\mathbf{1}_L}(\xi)| = O(1/\xi_1)$ because of the jump discontinuity in x_1 and smoothness in x_2 of $g\mathbf{1}_L$, so an elementary estimate with Taylor's theorem shows that $\lambda_1 > \lambda_2$. ■

Conversely, by taking $\tilde{g}(\xi)$ to be a smooth, radial, compactly supported function, we see that for any x^0 outside of the left half-plane L, and ϵ small enough, the dual local autocovariance matrix $E_{\epsilon,\tilde{g}}(\mathbf{1}_L; x^0)$ will be zero, since $\tilde{g}_\epsilon \mathbf{1}_L = 0$. Moreover, for each point x^0 in the interior of L, the matrix is a constant times the identity for all sufficiently small $\epsilon > 0$.

The eigenvector $(1,0)$ of the larger eigenvalue of the dual local autocovariance matrix at $(0,0)$ indicates the normal direction to the edge at $\{x_1 = 0\}$. It is easy to see that we will obtain the same result for every other edge point $(0, x_2)$, since we can translate $\mathbf{1}_L$ by such a vector without changing it or $|\widehat{g_\epsilon \mathbf{1}_L}|$. Likewise, if $R = \{(x_1, x_2) : x_1 \geq 0\}$ is the right half-plane, and $\mathbf{1}_R$ is its characteristic function, then $E_{\epsilon,g}(\mathbf{1}_R; x^0) = E_{\epsilon,g}(\mathbf{1}_L; x^0)$. That is because $\mathbf{1}_R(x) = \mathbf{1}_L(-x)$, and the dual local autocovariance matrix of $\mathbf{1}_L$ at 0 is preserved under this coordinate change. Other rotations and translations also give simple transformations:

Proposition 2.1. *Let T be a translation in $\mathbf{R}^2$ by x^0: $f \circ T(x) = f(x + x^0)$. Then $E_{\epsilon,g}(f; x^0) = E_{\epsilon,g}(f \circ T; 0)$.*

Proposition 2.2. *Let U be a rotation about x^0 in $\mathbf{R}^2$. Then $E_{\epsilon,g}(f \circ U; x^0) = U \circ E_{\epsilon,g}(f; x^0) \circ U^{-1}$.*

Of main interest to us is that the normal direction to an edge is an eigenvector of the larger eigenvalue of the dual local autocovariance matrix at an edge point, and that off edges the eigenvalues are the same. This gives us our first main result:

Theorem 2.1. *Suppose $H = \{x \in \mathbf{R}^2 : \nu \cdot x \leq \beta\}$ is a given half-plane, defined by the nonzero normal vector $\nu \in \mathbf{R}^2$ and some constant $\beta \in \mathbf{R}$. Let $\mathbf{1}_H$ be its characteristic function.*

1. *A point $x^0 \in \mathbf{R}^2$ belongs to ∂H if and only if for every smooth, radial nonzero function $g : \mathbf{R}^2 \to \mathbf{R}$, the matrix $E_{\epsilon,g}(\mathbf{1}_H; x^0)$ has distinct positive eigenvalues, and then the normal vector ν will be an eigenvector of the larger eigenvalue.*
2. *A point $x^0 \in \mathbf{R}^2$ is in the complement of ∂H if there is some smooth, nonzero but compactly-supported radial function $\tilde{g} : \mathbf{R}^2 \to \mathbf{R}$, and some $\epsilon > 0$, such that the matrix $E_{\epsilon,\tilde{g}}(\mathbf{1}_H; x^0)$ is a multiple of the identity.*

2.2. Domains with smooth boundary

Our calculations for the characteristic function of a half-plane also apply to the characteristic function $\mathbf{1}_D$ of a domain D with smooth boundary:

Proposition 2.3. *Let $D \subset \mathbf{R}^2$ be a domain with a smooth boundary, and fix $x^0 \in \partial D$. Fix a smooth radial function $g : \mathbf{R}^2 \to \mathbf{R}$, centered at 0 and supported in $B = B(0, 1)$. Let H be either half-plane defined by the line tangent to ∂D at x^0. Then for any matrix norm $\| \cdot \|$,*

$$\left\| E_{\epsilon,g}(\mathbf{1}_H; x^0) - E_{\epsilon,g}(\mathbf{1}_D; x^0) \right\| = O(\epsilon).$$

Proof. It will suffice to prove the estimate for the coefficients, since all norms are equivalent in the finite-dimensional space of 2×2 matrices. Also, note that if H, K are the two half-planes defined by the tangent line, then $E_{\epsilon,g}(\mathbf{1}_H; x^0) = E_{\epsilon,g}(\mathbf{1}_K; x^0)$ by Proposition 2.2. Hence it suffices to prove the result for one of them.

We may assume by Proposition 2.1 that x^0 is the origin. Then g_ϵ is supported in $B(0, \epsilon)$, so for one of the half-planes H defined by the line tangent to ∂D at x^0, we have $\left\|\widehat{g_\epsilon \mathbf{1}_H} - \widehat{g_\epsilon \mathbf{1}_D}\right\|_\infty \leq \|g_\epsilon \mathbf{1}_H - g_\epsilon \mathbf{1}_D\|_1 = O(\epsilon^3)$. Therefore

$$\left|E_{\epsilon,g}(\mathbf{1}_H; x^0)_{ij} - E_{\epsilon,g}(\mathbf{1}_D; x^0)_{ij}\right| \leq O(\epsilon^3) \int_{B(0,1/\epsilon)} |\xi_i||\xi_j| \int_{\mathbf{R}^2} |g_\epsilon(x)|\, dx d\xi = O(\epsilon).$$

■

A standard argument extends nearly the same estimate to a larger class of functions g:

Proposition 2.4. *Let $D \subset \mathbf{R}^2$ be a domain with a smooth boundary, and fix $x^0 \in \partial D$. Fix a Schwartz function $g : \mathbf{R}^2 \to \mathbf{R}$, radial about 0. Let H be either half-plane defined by the line tangent to ∂D at x^0. Then*

$$\left\|E_{\epsilon,g}(\mathbf{1}_H; x^0) - E_{\epsilon,g}(\mathbf{1}_D; x^0)\right\| = O(\epsilon^{1-\delta}),$$

as $\epsilon \to 0$, for every $\delta > 0$.

Proof. Again we assume without loss of generality that x^0 is the origin. Given any $\delta > 0$, let $\mu(\epsilon) = \epsilon^{-\delta/3}$. Then $\lim_{\epsilon\to 0} \mu(\epsilon) = \infty$, and $\epsilon^3\mu^3(\epsilon) = \mu^{3-9/\delta}$ is just a power of μ, so since g is rapidly decreasing,

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon^3\mu^3(\epsilon)} \int_{|x|\geq\mu(\epsilon)} g(x)\, dx = 0.$$

In particular, this gives a bound on the integral of g:

$$\int_{|x|\geq\mu(\epsilon)} g(x)\, dx = O(\epsilon^3\mu^3(\epsilon)). \tag{2}$$

But also, for one of the half-planes H the areas of $(H \setminus D) \cap B(x^0, r)$ and $(D \setminus H) \cap B(x^0, r)$ are both $O(r^3)$ as $r \to 0$. Hence, we obtain an estimate in two equivalent parts:

$$\begin{aligned}\left|\widehat{g_\epsilon \mathbf{1}_H}(\xi) - \widehat{g_\epsilon \mathbf{1}_D}(\xi)\right| &\leq \int_{\mathbf{R}^2} g_\epsilon(x)|\mathbf{1}_H(x) - \mathbf{1}_D(x)|\, dx \\ &= \int_{|x|\geq\epsilon\mu(\epsilon)} + \int_{|x|\leq\epsilon\mu(\epsilon)} \;=\; O\left(\epsilon^3\mu^3(\epsilon)\right) \;=\; O(\epsilon^{3-\delta}),\end{aligned}$$

because of our choice of μ. We may now reuse the last part of the proof of Proposition 2.3 to obtain the result:

$$|E_{\epsilon,g}(\mathbf{1}_H; x^0)_{ij} - E_{\epsilon,g}(\mathbf{1}_D; x^0)_{ij}| = O(\epsilon^{1-\delta}),$$

for $i, j \in \{1, 2\}$. By Proposition 2.2, this holds for the other half-plane as well. ■

We may now state our second main result:

Theorem 2.2. *Suppose $D \subset \mathbf{R}^2$ is a domain with a smooth boundary, $x^0 \in \partial D$, and $g : \mathbf{R}^2 \to \mathbf{R}$ is a nonzero radial Schwartz function. Let H be either half-plane* *defined by the line tangent to ∂D at x^0. If we denote the eigenvalues of $E_{\epsilon,g}(\mathbf{1}_D; x^0)$ by $\lambda_1(\epsilon)$ and $\lambda_2(\epsilon)$, then* $\liminf_{\epsilon\to 0+} |\lambda_1(\epsilon) - \lambda_2(\epsilon)| > 0$.

Proof. By Lemma 2.1, the eigenvalues of $E_{\epsilon,g}(\mathbf{1}_H; x^0)$ satisfy $\lambda_1 > \lambda_2 > 0$, and do not depend on ϵ. The result follows from Proposition 2.4. ■

The boundary smoothness assumption in Proposition 2.3 and Proposition 2.4 is not crucial. We can prove similar results for rougher ∂D, using the weaker Morrey-Campanato regularity assumption [4] in generalized form [1]. This leads to the following result:

Theorem 2.3. *Suppose D is a domain in $\mathbf{R}^2$ such that ∂D is of generalized Morrey-Campanato class $L(\rho, \infty, 1)$ for some function $\rho(\epsilon) = o(\epsilon)$, as $\epsilon \to 0$. Fix $x^0 \in \partial D$, and suppose further that for every $\epsilon > 0$, there exists a line ℓ_ϵ that realizes the infimum of $\|\partial D - \ell_\epsilon\|_{L(\rho,\infty,1)}$ over the ball $B(x^0, \epsilon)$. Choose a smooth, radial function $g : \mathbf{R}^2 \to \mathbf{R}$ with compact support in $B = B(0, 1)$. Let $H = H_\epsilon$ be either half-plane defined by ℓ_ϵ. Then*

$$\lim_{\epsilon\to 0} \left\| E_{\epsilon,g}(\mathbf{1}_H; x^0) - E_{\epsilon,g}(\mathbf{1}_D; x^0) \right\| = 0.$$

Proof. By the definition of the space $L(\rho, \infty, 1)$, for the family of half-planes H_ϵ chosen as above, we have that:

$$\|g_\epsilon(\mathbf{1}_H - \mathbf{1}_D)\|_1 = O(\epsilon\, \rho(\epsilon)).$$

Repeating the calculations from the proof of Proposition 2.1, we obtain:

$$\left\| E_{\epsilon,g}(\mathbf{1}_H; x^0) - E_{\epsilon,g}(\mathbf{1}_D; x^0) \right\| = O(\epsilon^{-1}\rho(\epsilon)).$$

The result follows from $\rho(\epsilon) = o(\epsilon)$. ■

If we allow g to be a radial Schwartz function, then a standard modification of this proof, along the lines of Proposition 2.4, yields the estimate

$$\left\| E_{\epsilon,g}(\mathbf{1}_H; x^0) - E_{\epsilon,g}(\mathbf{1}_D; x^0) \right\| = O(\rho(\epsilon)\epsilon^{-1-\delta})$$

as $\epsilon \to 0$, for every $\delta > 0$.

An example of a domain with Morrey-Campanato regular boundary is shown in Figure 1.

2.3. Functions with tangent planes

The dual local autocovariance matrix of a function f, at a point along a Morrey-Campanato curve of discontinuities, has distinct eigenvalues. Now, we show that at differentiable points the eigenvalues are equal.

Proposition 2.5. *Let $A : R^2 \to \mathbf{R}$ be an affine function: $A(x) = a \cdot x + b$ for some constants $a \in \mathbf{R}^2, b \in \mathbf{R}$. Let $g : \mathbf{R}^2 \to \mathbf{R}$ be a smooth radial function*

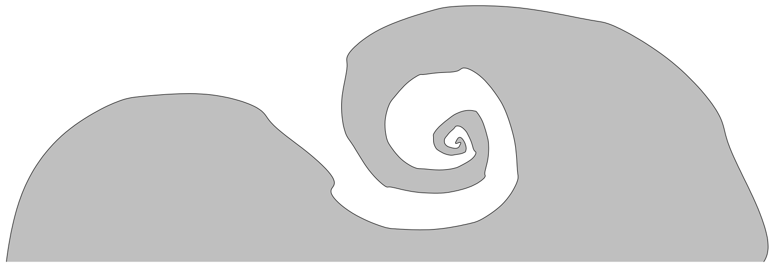

FIG. 1. Domain with Morrey-Campanato regular boundary. At the point of the spiral, approximate tangents exist at all scales ϵ, but do not converge as $\epsilon \to 0$.

supported in $B = B(0,1)$. Denote by $\lambda_1(\epsilon), \lambda_2(\epsilon)$ the two eigenvalues of the dual local autocovariance matrix $E_{\epsilon,g}(A;0)$. Then:

$$\lim_{\epsilon \to 0+} [\lambda_1(\epsilon) - \lambda_2(\epsilon)] = 0.$$

Proof. Writing ∇ for the normalized gradient $\frac{-1}{2\pi i}(\frac{\partial}{\partial \xi_1}, \frac{\partial}{\partial \xi_2})$, we get:

$$E_{\epsilon,g}(A;0)_{ij} = \int_{B(0,1/\epsilon)} \xi_i \xi_j \left| \widehat{Ag_\epsilon} \right|^2 d\xi = \int_{B(0,1/\epsilon)} \xi_i \xi_j \left| a \cdot \nabla \hat{g}_\epsilon + b\hat{g}_\epsilon \right|^2 d\xi.$$

Now $\left| a \cdot \nabla \hat{g}_\epsilon + b\hat{g}_\epsilon \right|^2 = \left| a \cdot \nabla \hat{g}_\epsilon \right|^2 + \left| b\hat{g}_\epsilon \right|^2 + 2\Re\left\{ (a \cdot \nabla \hat{g}_\epsilon)(b\hat{g}_\epsilon) \right\}$. We may assume, after any needed translation, that $x^0 = 0$. Then the pieces evaluate as follows:

- $\hat{g}_\epsilon(\xi) = \epsilon^2 \hat{g}(\epsilon\xi)$, so

$$\int_{B(0,1/\epsilon)} \xi_i \xi_j |b\hat{g}_\epsilon(\xi)|^2 \, d\xi = \int_B \eta_i \eta_j |b\hat{g}(\eta)|^2 \, d\eta = C\delta_{ij},$$

by the symmetry of g and thus $\hat{g}$. Here $C = |b|^2 \int_B \eta_1^2 |\hat{g}|^2 = |b|^2 \int_B \eta_2^2 |\hat{g}|^2$ is a nonnegative constant depending on g, but independent of ϵ.

- $\nabla \hat{g}_\epsilon(\xi) = \epsilon^3 \nabla \hat{g}(\epsilon\xi)$, so

$$\int_{B(0,1/\epsilon)} \xi_i \xi_j |a \cdot \nabla \hat{g}_\epsilon(\xi)|^2 \, d\xi = \epsilon^2 \int_B \eta_i \eta_j |a \cdot \nabla \hat{g}(\eta)|^2 \, d\eta = O(\epsilon^2).$$

- ξ_1, ξ_2, and ϵ are all real, so

$$\int_{B(0,1/\epsilon)} \xi_i \xi_j \Re\{a \cdot \nabla \hat{g}_\epsilon(\xi) b\hat{g}_\epsilon(\xi)\} \, d\xi = \epsilon \Re\left\{ \int_B \eta_i \eta_j a \cdot \nabla \hat{g}(\eta) b\hat{g}(\eta) \, d\eta \right\} = O(\epsilon).$$

In each case, the substitution $\xi \leftarrow \epsilon\eta$ rescales the domain of integration to $B = B(0,1)$. Thus

$$E_{\epsilon,g}(A;x^0)_{ij} = C\delta_{ij} + O(\epsilon), \qquad \text{as } \epsilon \to 0,$$

from which the result follows. ∎

PROPOSITION 2.6. *Suppose that $f : \mathbf{R}^2 \to \mathbf{R}$ is differentiable at x^0, and let A be the affine function tangent to f at x^0. Fix a smooth function g of compact support. Then $\left\|E_{\epsilon,g}(f;x^0) - E_{\epsilon,g}(A;x^0)\right\| = o(\epsilon)$, as $\epsilon \to 0$. If, in addition, f is continuously differentiable in a neighborhood of x^0, then in fact*

$$\left\|E_{\epsilon,g}(f;x^0) - E_{\epsilon,g}(A;x^0)\right\| = O(\epsilon^2), \qquad \text{as } \epsilon \to 0.$$

Proof. We may write $f = A + \omega$, where $A(x) = a \cdot x + b$ and $\omega(x) = o(|x - x^0|)$ in some neighborhood of x^0. We compare the dual local autocovariance matrices of f and A:

$$E_{\epsilon,g}(f;x^0)_{ij} - E_{\epsilon,g}(A;x^0)_{ij} = \int_{B(0,1/\epsilon)} \xi_i \xi_j \left(|\widehat{\omega g_\epsilon}|^2 + 2\Re\{\widehat{\omega g_\epsilon}\overline{\widehat{Ag_\epsilon}}\} \right) d\xi.$$

Now $|\omega(x)| = o(\epsilon)$ on the support of g_ϵ, so $|\widehat{\omega g_\epsilon}(\xi)| = o(\epsilon^3)\|g\|_1$, as $\epsilon \to 0$, uniformly in ξ. In addition, $|\Re\{\widehat{Ag_\epsilon}(\xi)\}| \le |\widehat{Ag_\epsilon}(\xi)| = O(\epsilon^2)\|Ag\|_1$, uniformly in ξ. Thus:

$$\left|E_{\epsilon,g}(f;x^0)_{ij} - E_{\epsilon,g}(A;x^0)_{ij}\right| = o(\epsilon^5) \int_{B(0,1/\epsilon)} |\xi_i \xi_j| \, d\xi = o(\epsilon) \int_B |\eta_i \eta_j| \, d\eta = o(\epsilon).$$

If f is continuously differentiable in a neighborhood of x^0, then $\omega(x) = O(|x - x^0|^2)$ near x^0 and the same argument results in the better estimate. ■

Combining Propositions 2.5 and 2.6 gives our third main result:

THEOREM 2.4. *Suppose that $f : \mathbf{R}^2 \to \mathbf{R}$ is differentiable at x^0. Then for any smooth radial function $g : \mathbf{R}^2 \to \mathbf{R}$ of compact support, the matrix $E_{\epsilon,g}(f;x^0)$ converges to a multiple of the 2×2 identity matrix, as $\epsilon \to 0$.*

The converse to Theorem 2.4 is false: even if $\lim_{\epsilon \to 0+} |E_{\epsilon,g}(f;x^0)|$ is a multiple of the identity, we cannot conclude that f is differentiable at x^0, or even continuous. Sufficient symmetry can masquerade as smoothness, as the following example shows. Let $f(x) = \mathbf{1}_+(x_1)\mathbf{1}_+(x_2)$, where $\mathbf{1}_+$ is the characteristic function of $\mathbf{R}^+$, and fix $g(x) = \exp(-\pi|x|^2)$ as before. Then, reusing some calculations from Lemma 2.1, we discover that $E_{\epsilon,g}(f;0)$ is a positive multiple of the identity, so $\lambda_1(\epsilon) = \lambda_2(\epsilon) = \lambda > 0$ for every ϵ, even though f is discontinuous at 0.

2.4. Higher dimensional theory

Dual local autocovariance matrices can be defined for functions $f : \mathbf{R}^p \to \mathbf{R}^d$, and provide a tool to study the geometry of complicated maps in high dimensions.

DEFINITION 2.2. Let $f : \mathbf{R}^p \to \mathbf{R}^d$ be a polynomially bounded, vector valued function, $f = (f_1, \ldots, f_d)$, and fix a point $x^0 \in \mathbf{R}^p$. Choose a radial function $g : \mathbf{R}^p \to \mathbf{R}^d$ in the Schwartz class, and let $\epsilon > 0$ be given. Then we may define the *dual local autocovariance matrix of f at x^0* to be the $p \times p$ matrix:

$$E_{\epsilon,g}(f;x^0)_{ij} = \int_{B(0,1/\epsilon)} \xi_i \xi_j \|\widehat{f \cdot g_\epsilon}(\xi)\|^2 \, d\xi,$$

for $i, j \in \{1, \ldots, p\}$, where $g_\epsilon(x) = g\left(\frac{x-x^0}{\epsilon}\right)$, as before.

This p-dimensional dual local autocovariance matrix preserves the properties of the two-dimensional example. In particular, we have the following immediate generalization of Theorem 2.4.

THEOREM 2.5. *Suppose that $f : \mathbf{R}^p \to \mathbf{R}^d$ is differentiable at x^0. Then for any smooth radial function $g : \mathbf{R}^p \to \mathbf{R}^d$ of compact support, the matrix $E_{\epsilon,g}(f; x^0)$ converges to a multiple of the $p \times p$ identity matrix, as $\epsilon \to 0$.*

Theorem 2.3 also generalizes. We may define a hyperplane $\ell = \{x \in \mathbf{R}^p : a \cdot x = b\}$ and half-space $H = \{x \in \mathbf{R}^p : a \cdot x \leq b\}$ by constants $a \in \mathbf{R}^p, b \in \mathbf{R}$, and we may denote by $\lambda_i(\epsilon)$ the i^{th} eigenvalue of the matrix $E_{\epsilon,g}(\mathbf{1}_H)$, for $g(x) = \exp(-\pi\|x\|^2)$, arranged in decreasing order. Lemma 2.1 generalizes to imply $\lambda_1(\epsilon) > \lambda_j(\epsilon)$ for all $j = 2, \ldots, p$, independent of ϵ.

THEOREM 2.6. *Suppose D is a domain in $\mathbf{R}^p$ such that ∂D is of generalized Morrey-Campanato class $L(\rho, \infty, 1)$ for some function $\rho(\epsilon) = o(\epsilon)$, as $\epsilon \to 0$. Fix $x^0 \in \partial D$, and suppose further that for every $\epsilon > 0$, there exists a hyperplane ℓ_ϵ that realizes the infimum of $\|\partial D - \ell_\epsilon\|_{L(\rho,\infty,1)}$ over the ball $B(x^0, \epsilon)$. Choose a smooth radial function $g : \mathbf{R}^p \to \mathbf{R}$ with compact support in $B = B(0, 1)$. Let $H = H_\epsilon$ be either half-space defined by ℓ_ϵ. Then*

$$\lim_{\epsilon \to 0} \left\| E_{\epsilon,g}(\mathbf{1}_H; x^0) - E_{\epsilon,g}(\mathbf{1}_D; x^0) \right\| = 0.$$

3. CASE STUDIES

3.1. Algorithm and implementation

We make some well-known approximations to calculate the autocovariance matrix in the discrete sampled case. Our normalization of the one-dimensional discrete Fourier transform on N real samples $\{f(n) : 0 \leq n < N\}$ is

$$\hat{f}(k) = \sum_{n=0}^{N-1} \exp\left(-2\pi i \frac{kn}{N}\right) f(n), \qquad k \in B_N = \left[-\frac{N}{2}, \frac{N}{2}\right].$$

If only the first $q \ll N$ samples of f are nonzero, then the sum reduces to the smaller range $\{0, 1, \ldots, q-1\}$. Thus, the squared absolute value of $\hat{f}(k)$, when f is real-valued, is

$$|\hat{f}(k)|^2 = \sum_{n=0}^{q-1} \sum_{n'=0}^{q-1} \exp\left(-2\pi i \frac{k(n-n')}{N}\right) f(n) f(n'), \qquad k \in B_N.$$

The r-th moment of $|\hat{f}(k)|^2$ is therefore

$$\sum_{k \in B_N} k^r |\hat{f}(k)|^2 \ = \ \sum_{k \in B_N} \sum_{n=0}^{q-1} \sum_{n'=0}^{q-1} k^r \exp\left(-2\pi i \frac{k(n-n')}{N}\right) f(n) f(n')$$

$$= \sum_{n=0}^{q-1}\sum_{n'=0}^{q-1} f(n)f(n') \sum_{k\in B_N} k^r \exp\left(-2\pi i \frac{k(n-n')}{N}\right).$$

The innermost sum in k is a function of the integer $n-n'$. Except for a factor of N^{r+1}, it is a Riemann approximation to the integral

$$\mu_r(n) = \int_{-\frac{1}{2}}^{\frac{1}{2}} x^r \exp(-2\pi i n x)\,dx,$$

evaluated at $n \leftarrow n-n'$, whose easily-computed values we shall use instead. Evidently $\mu_0(n)=1$ if $n=0$, but is zero otherwise, while

$$\mu_1(n) = \begin{cases} 0, & \text{if } n=0,\\ i\frac{(-1)^n}{2\pi n}, & \text{otherwise;}\end{cases} \qquad \mu_2(n) = \begin{cases} 1/12, & \text{if } n=0,\\ \frac{(-1)^n}{2\pi^2 n^2}, & \text{otherwise.}\end{cases}$$

We can apply the above results to analyze an image, which for our purposes will be a real-valued function supported on the rectangle $[0,M]\times[0,N]\subset\mathbf{R}^2$, sampled on a regular grid with grid point coordinates $\{(m,n): 0\le m<M;\ 0\le n<N\}$. We will use a bump function supported on small subrectangles of size $p\times q$, rather than a dilated radial function, for g_ϵ. Translations of f have no effect on $|\hat{f}|^2$, so we may assume that the localized portion of the image has been translated to the subgrid $\{(m,n): 0\le m<p;\ 0\le n<q\}$. The dual local autocovariance matrix may then be computed as follows:

$$\begin{aligned} E_{11} &= \sum_{m,m',n,n'} f(m,n)f(m',n')\mu_2(m-m')\mu_0(n-n') \qquad (3)\\ &= \sum_{m=0}^{p-1}\sum_{m'=0}^{p-1}\sum_{n=0}^{q-1} f(m,n)f(m',n)\mu_2(m-m');\\ E_{22} &= \sum_{m=0}^{p-1}\sum_{n=0}^{q-1}\sum_{n'=0}^{q-1} f(m,n)f(m,n')\mu_2(n-n'); \qquad (4)\\ E_{12} = E_{21} &= \sum_{m=0}^{p-1}\sum_{m'=0}^{p-1}\sum_{n=0}^{q-1}\sum_{n'=0}^{q-1} f(m,n)f(m',n')\mu_1(m-m')\mu_1(n-n'). \qquad (5)\end{aligned}$$

Around each grid point x^0 of the image, we perform the following steps:

Localization. Extract the samples on the square subgrid $x^0+[-\epsilon,\epsilon]\times[-\epsilon,\epsilon]$, where a small positive integer plays the role of ϵ in Definition 2.1. Then $p=q=2\epsilon+1$. The sample at x^0 becomes the sample at 0 in the $(2\epsilon+1)\times(2\epsilon+1)$ extracted subgrid $[-\epsilon,\epsilon]^2$. The sample at x^0+y is multiplied by the Gaussian bump function $\exp(-\pi|y|^2/\epsilon^2)$ and stored at y in the subgrid. This costs $O(\epsilon^2)$ operations per pixel.

If x^0 is within ϵ grid points of the boundary, then we simply pad any missing samples in the subgrid with zeros. For definiteness, we chose $\epsilon=3$ to prepare our case studies.

Dual Autocovariance. Compute the 2×2 matrix $E=(E_{ij})$ using Equations 3,4, and 5. This always yields a real-valued, symmetric, positive semidefinite matrix.

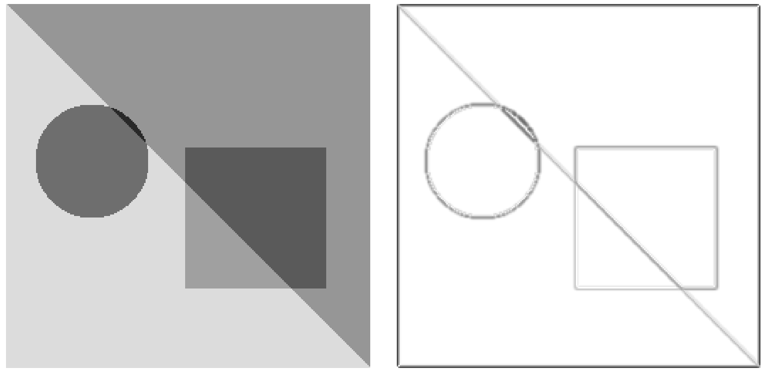

FIG. 2. Geometrical figures: image, and edginess.

The computational complexity of the quadruple sum in Equation 5 dominates the triple sums of Equations 3 and 4, so this step costs $O(\epsilon^4)$ operations per pixel.

Eigenvalues. For symmetric 2×2 matrices E, the exact formula for eigenvalues is:

$$\lambda = \frac{1}{2}\left(E_{11} + E_{22} \pm \sqrt{(E_{11} - E_{22})^2 + 4E_{12}^2}\right), \tag{6}$$

where we take $+$ for λ_1 and $-$ for λ_2. These will satisfy $\lambda_1 \geq \lambda_2 \geq 0$, so in particular, if $\lambda_2 > 0$ we always have $\lambda_1/\lambda_2 \geq 1$. The greater the relative difference, the greater the edginess.

Edginess. We define this to be the ratio λ_1/λ_2. We actually compute the bounded reciprocal $\lambda_2/\lambda_1 \in [0, 1]$, amplified to fill the grayscale range of a write-black display device. That way, the darkest marks indicate the greatest edginess.

The Dual Autocovariance step dominates the computational complexity. It is therefore $O(p^2q^2)$ operations per pixel, if we localize to subgrids of $p \times q$ points.

3.2. Example images

We prepared five examples by the algorithm described above, localizing with Gaussian bumps restricted to 7×7 subgrids centered at x^0.

The geometrical figures are piecewise constant functions with jump discontinuities along various rectifiable, mostly smooth curves. The fingerprint image was obtained from NIST; it is part of its compliance test suite for the FBI's WSQ compression standard. "Lena" is the famous image from [7]. "Cone" is a synthetic ray-traced image provided by Craig Kolb. "Truck," provided by Peng Li, is one frame from a video. All five are available from the ACHA Software Distribution Web Site.

REFERENCES

1. W. K. Czaja. *Applications of Local Autocovariance Matrices.* PhD thesis, Washington University, Saint Louis, Missouri, 2000.
2. R. Duda and P. Hart. Use of the Hough transform to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972.
3. A. Gelb and E. Tadmor. Detection of edges in spectral data. *Applied and Computational Harmonic Analysis*, 7(1):101–135, 1995.

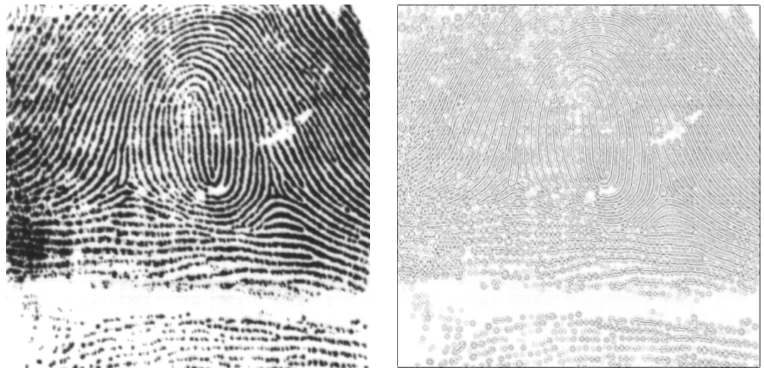

FIG. 3. Fingerprint: image, and edginess.

FIG. 4. Lena: image, and edginess.

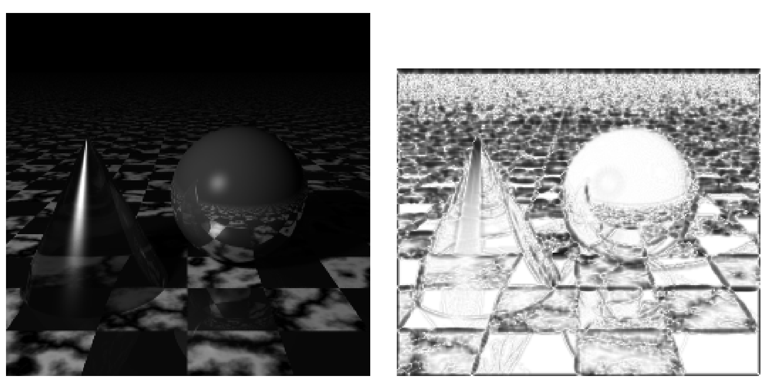

FIG. 5. Cone: image, and edginess.

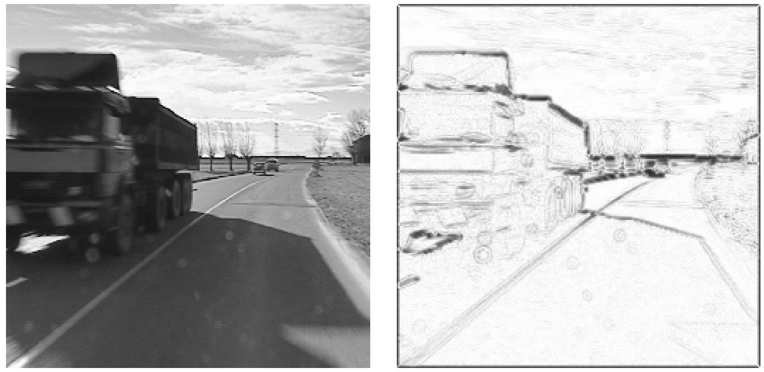

FIG. 6. Truck: image, and edginess.

4. S. Janson, M. Taibleson, and G. Weiss. Elementary characterizations of the Morrey-Campanato spaces. In G. Weiss, G. Mauceri, and F. Ricci, editors, *Proceedings of the July, 1982 Harmonic Analysis Conference Held in Cortona, Italy*, volume 992 of *Lecture Notes in Mathematics*, pages 101–114, New York, 1983. Springer-Verlag.
5. L. Kurz and M. H. Benteftifa. *Analysis of Variance in Statistical Image Processing.* Cambridge University Press, Cambridge, England, 1997.
6. S. G. Mallat and S. Zhong. Wavelet transform maxima and multiscale edges. In M. B. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Raphael, editors, *Wavelets and Their Applications*, pages 67–104. Jones and Bartlett, Boston, 1992.
7. L. Sjooblum. Miss November. Playboy Magazine, Playboy Enterprises, Inc., 1972.
8. M. V. Wickerhauser. Large-rank approximate principal component analysis with wavelets for signal feature discrimination and the inversion of complicated maps. *Journal of Chemical Information and Computer Science*, 34(5):1036–1046, September/October 1994.