Natural demodulation of two-dimensional fringe patterns. I. General background of the spiral phase quadrature transform

Kieran G. Larkin

Canon Information Systems Research Australia, 1 Thomas Holt Drive, North Ryde, Sydney, NSW 2113, Australia, and Department of Physical Optics, School of Physics, The University of Sydney, Sydney, NSW 2006, Australia

Donald J. Bone

Advanced Image Research, P.O. Box 1001, Dickson ACT 2602, Australia

Michael A. Oldfield

Canon Information Systems Research Australia Pty Ltd, 1 Thomas Holt Drive, North Ryde, Sydney, NSW 2113, Australia

(Received August 24, 2000; revised manuscript received January 16, 2001; accepted January 22, 2001)

It is widely believed, in the areas of optics, image analysis, and visual perception, that the Hilbert transform does not extend naturally and isotropically beyond one dimension. In some areas of image analysis, this belief has restricted the application of the analytic signal concept to multiple dimensions. We show that, contrary to this view, there is a natural, isotropic, and elegant extension. We develop a novel two-dimensional transform in terms of two multiplicative operators: a spiral phase spectral (Fourier) operator and an orientational phase spatial operator. Combining the two operators results in a meaningful two-dimensional quadrature (or Hilbert) transform. The new transform is applied to the problem of closed fringe pattern demodulation in two dimensions, resulting in a direct solution. The new transform has connections with the Riesz transform of classical harmonic analysis. We consider these connections, as well as others such as the propagation of optical phase singularities and the reconstruction of geomagnetic fields. © 2001 Optical Society of America

OCIS codes: 100.2650, 100.2960, 100.5070, 100.5090, 120.3180, 120.5050.

1. INTRODUCTION: HISTORICAL SURVEY

Our work on the natural demodulation of two-dimensional (2-D) fringe patterns will be presented in two parts. In this, Paper I, we present a background to the heuristic derivation of the spiral phase quadrature transform and some simulations. In Paper II (this issue) a mathematical basis for the validity and the accuracy of the spiral phase quadrature transform is investigated. The work presented is the culmination of several years of investigating mathematical methods for the demodulation of human fingerprints and other naturally occurring fringe patterns. Although the presentation concentrates on the logical development of the isotropic quadrature operator, it glosses over the rather convoluted experimental development of the technique. The discovery that our method of isotropic demodulation is closely related to the Riesz transform occurred after the technique was working effectively as an image processing operation. Similarly, the mathematical justification for the heuristic method was developed only after extensive experimental testing on real and simulated fringe patterns. The chronicle might have been different if there were less conflicting information on the Hilbert transform (HT) in two dimensions.

To help understand the confusion regarding the extension of the HT beyond one dimension, a brief historical survey of the literature is useful. The concept of an analytic (or holomorphic) signal was introduced to communication theory by Gabor[1] in 1947 for one-dimensional (1-D) signals. An analytic signal consists of two parts: The real part is the base signal, and the imaginary (or quadrature) part is the HT of the real part. The theory of analytic signals naturally underpins many modern concepts of signal analysis such as amplitude and frequency (AM–FM) demodulation, spectral analysis, instantaneous frequency, interferometry, and radar. Unfortunately, the concept has not, apparently, extended naturally beyond one dimension without implying a preferred direction. Consequently, a number of avowedly *ad hoc* definitions of the 2-D HT have been proposed[2–7] with varying degrees of directionality. Typical definitions have half-plane symmetry,[5] quadrant-based symmetry,[8,9] or rotated half-plane symmetry.[10] A recent development is the idea of extending the complex analysis of the Fourier transform (FT) to hypercomplex numbers. The concept has been called "hyper-complex signal representation" by Bülow and Sommer[11,12] and potentially allows an unambiguous definition of the analytic image in two dimensions. Unfortunately and surprisingly, the published definition has a degree of directionality that is apparent in the demodulated envelope patterns.[11] A similar idea using quaternions and even octonions for multidimensional signals

0740-3232/2001/081862-09$15.00 © 2001 Optical Society of America

was proposed by Craig[13] in 1996. The quaternionic approach allows several possible definitions but introduces additional phases into the definition of the analytic image.

In the area of phase retrieval, the concept of analyticity is central to the understanding of multidimensional bandlimited signals.[14] Interestingly, the mathematical development of a complex function of several complex variables (and the associated Cauchy–Riemann equations) leads, in this case, to a nonisotropic interpretation of the HT relations in the two real-space variables.[15,16] However, an alternative definition of the multidimensional Cauchy–Riemann conditions leads to isotropic equations.[17] In an isotropic system it is not clear why there should be a preferred direction. Sometimes, anisotropic definitions are justified by the symmetry of the problem. For example, images obtained by differential interference contrast microscopy have one direction related to the differential shear, so the application of a directional multidimensional HT may be appropriate.[18] Similarly, in three-dimensional (3-D) white-light interferometry the HT relation applies to just one coordinate.[19]

Little known to many researchers in signal processing, the theory of the HT extended to n dimensions (n real variables) has been in development since the 1920s by pure mathematicians working in an area known as the harmonic analysis of singular integrals. Following Hilbert's lead,[20] Riesz[21] proposed "fonctions conjuguées," or conjugate functions, as extensions to the HT. Subsequently, independent work by Tricomi[22] and Giraud[23] developed the same idea. More recently, the works of Mikhlin,[24,25] followed soon after by that of Calderon and Zygmund,[26,27] have proved the existence and the convergence of the associated integral operators.[28] Another approach to the problem by the generalization of the Cauchy–Riemann conditions to higher dimensions was undertaken by Fulton and Rainich.[17] Readers wishing to follow the rather circuitous development of the n-dimensional Riesz transform (RT) (as the n-dimensional analog of the HT is now known) are advised to start with the textbooks by Stein[29] and Mikhlin[25] and the paper by Carberry.[30] The main complication with the RT is that it is an n-vector for an n-dimensional scalar signal, and the corresponding analytic signal is an $(n + 1)$-vector. In 2-D image processing the resulting signal is a three-vector and cannot be displayed as a complex image.

Of all the applied sciences, geophysics has been perhaps the most successful in finding possible definitions of the 2-D HT over the years.[13,31–34] Indeed, the definitions of Nabighian[32] and Craig[13] developed for geomagnetic field analysis coincide with the definition of the RT, although they do not explicitly refer to the RT in their work.

A number of researchers have claimed that a true 2-D HT can be considered difficult[35] or impossible.[6,36–39] The difficulty is based on the perceived problem of extending the 1-D signum function, central to the 1-D HT, to a 2-D signum function. The spiral phase formalism for the 2-D HT developed in this paper marks a conceptual change from the incumbent linear signum function to a revolutionary signum function. To demonstrate the power of the new formalism, we use an otherwise intractable problem in fringe pattern analysis, which becomes almost trivial because the usual difficulty in linearly separating spectral zones does not occur.

Two recent publications have touched on the idea of an isotropic HT. The first[40] (written in German, but our translation is available to interested researchers) explicitly uses the 2-D RT to enhance digital images. The second[41] considers a "radial HT" for digital image enhancement but in the context of an optical spiral phase filter implemented with a spatial light modulator. Neither publication discusses the rather significant quadrature effects of the transform. In this publication we shall concentrate upon the remarkable phase- and quadrature-related effects rather than the intensity or magnitude effects seen in digital images.

2. BACKGROUND: ONE-DIMENSIONAL HILBERT TRANSFORM

Details of the conventional HT $\mathcal{H}$ and analytic signal are well described in Bracewell's classic textbook.[42] Perhaps the most important property for signal processing is that it transforms all cosine components in a function of x to sines and vice versa, regardless of scale factor λ:

$$-\sin(\lambda x) = \mathcal{H}\{\cos(\lambda x)\} \qquad \text{for } \lambda > 0, \tag{1}$$

$$\cos(\lambda x) = \mathcal{H}\{\sin(\lambda x)\}. \tag{2}$$

In many cases the Fourier (or spectral) description of the HT is informative:

$$\hat{f}(x) = \mathcal{H}\{f(x)\} \tag{3}$$

defines the HT of a real function f, and

$$p(x) = f(x) - i\hat{f}(x) = |b(x)|\exp[i\psi(x)] \tag{4}$$

defines the corresponding complex analytic signal.

The FT operator $\mathcal{F}$, operating on g, is given by

$$G(u) = \int_{-\infty}^{+\infty} g(x)\exp(-2\pi i u x)\mathrm{d}x = \mathcal{F}\{g(x)\}, \tag{5}$$

whereas the FT of the HT of g is given by

$$i\,\mathrm{sign}(u)G(u) = \mathcal{F}\{\hat{g}(x)\}. \tag{6}$$

In other words, the FT of the HT of g is the FT of g multiplied by an imaginary signum function (see Fig. 1). Note that g and $\hat{g}$ are real functions. Many attempts to

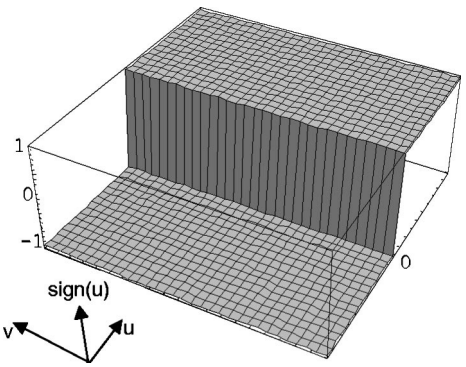

Fig. 1. Half-plane signum function plotted as a function of frequency coordinates (u, v).

extend the HT are based on extending the signum function to two dimensions. Conventionally, these are products of 1-D functions, which result in the half-plane and quadrant signum functions. Such products are highly anisotropic owing to the directional line discontinuities, culminating in the familiar anisotropic definitions. A more promising approach, conceptually, is to maintain the point discontinuity of the 1-D signum function in higher dimensions. Clearly, a point is a nondirectional discontinuity in two or more dimensions. The transforms of several authors actually have a point discontinuity[13,32]; however, the output is not a scalar for a scalar input but is either a two-vector or a quaternion. It is not immediately clear how to interpret the output in such cases, and this may explain to some extent why these methods have not been endorsed more generally.

3. TWO-DIMENSIONAL QUADRATURE FUNCTIONS

We are quite clear about what we demand of a 2-D quadrature transform, even if a 2-D HT is difficult to agree upon. A classical problem in fringe analysis illustrates the requirements rather well. A typical 2-D fringe pattern has the following form:

$$f(x,y) = a(x,y) + b(x,y)\cos[\psi(x,y)]. \tag{7}$$

Typically, the offset and modulation terms, a and b, respectively, are slowly and smoothly varying functions. The phase function ψ is also smoothly varying, but the combined effect is a rapidly oscillating function f. The objective of fringe pattern analysis (also the objective of AM–FM signal demodulation in general) is to extract the amplitude and phase functions, $b(x,y)$ and $\psi(x,y)$, respectively. One of the most powerful methods—known as the Fourier transform method (FTM)—was originally developed for one dimension[43,44] and subsequently extended to two dimensions.[45] The two main complications in two dimensions are that the FTM cannot separate the overlapping spectral components of closed-curve fringes and that there are local and global ambiguities in the sign ($\pm$) of the output quadrature estimate.[46] Until now, no direct methods have been able to surmount these obstacles (note that indirect methods using either computationally intensive optimization algorithms[47] or extensive manual intervention can succeed). We can identify two key points in the development of a direct 2-D quadrature method. The first is, in accordance with conventional belief, the definition of a suitable 2-D signum function in the spectral domain. The second is that a 2-D signum function alone is unable to ensure that the output is real (for real input), of the correct polarity, and direction insensitive. To do this, we propose a second operation purely in the spatial domain.

The full 2-D Fourier domain analysis of our proposed operator is presented in Paper II. Our proposed 2-D signum function is defined simply as a pure spiral phase function in spatial frequency space (u,v):

$$S(u,v) = \frac{u + iv}{\sqrt{u^2 + v^2}} = \exp[i\phi(u,v)]. \tag{8}$$

Here the phase ϕ is the polar angle in frequency space. The spiral phase function has the curious property that any section through the origin is a signum function. The major influence in the conceptual and mathematical development of our spiral phase formalism has been the research on optical vortices. Nye and Berry[48] first showed that edge dislocations can exist in 3-D waves and that the phase around the edge resembles a vortex. There are deep connections mainly related to the Fourier property of far-field diffraction patterns.[48–51] Spiral phase plates or holograms in the Fourier plane are analogous to the Fourier multipliers of singular integrals. Another connection is that the spiral phase discontinuity (in the guise of a residue) is central to the theory of phase unwrapping in two dimensions, as comprehensively described by Ghiglia and Pritt.[52] The phase-unwrapping connection occurs again in the definition of orientation and is discussed further in Paper II. It transpires that the phase spiral is also consistent with definitions of the 2-D RT represented by a complex quantity instead of a two-vector. Figure 2 shows a representation of the principal value (p.v.) of the spectral polar angle $\phi(u,v)$. The 2π discontinuity in the phase p.v. $\phi(u,v)$ is unimportant because the complex exponential in Eq. (8) is continuous everywhere (except the origin). Our reason for using the spiral phase function S is that it has the following properties:

1. It has odd radial symmetry, $-S(u,v) = S(-u,-v)$, so that it converts odd radial functions to even, and even radial functions to odd.
2. It contains only a single point discontinuity (maintaining circular symmetry).
3. There is no radial variation of magnitude or phase with radius, and the magnitude is unity, hence ensuring scale invariance.
4. The relative angular variation is constant, so that it has uniform rotational properties.

This spiral phase Fourier multiplier is applied to $g(x,y)$, a fringe pattern with its offset removed[53]:

$$g(x,y) = f(x,y) - a(x,y) = b(x,y)\cos[\psi(x,y)]. \tag{9}$$

Hence the ideal quadrature function (assuming suitably band-limited amplitude and phase[54]) would be

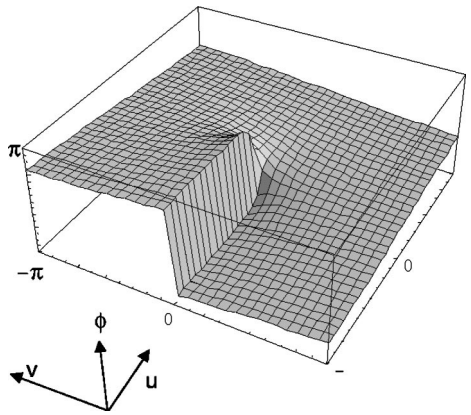

Fig. 2. Spiral phase "signum" function exponent ϕ. The principal value of the complex exponent $\phi(u,v)$ is shown in the range $\pm\pi$.

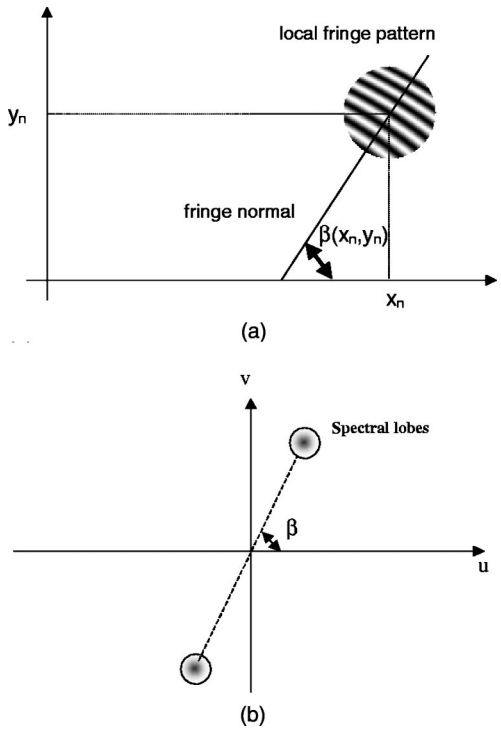

Fig. 3. (a) Definition of local orientation angle $\beta(x, y)$ for a locally simple fringe pattern. Each point in the fringe pattern has a well-defined orientation angle. (b) Spectral sidelobes related to the local fringe pattern. The lobes are located at a polar angle equal to the fringe normal angle.

$$\hat{g}(x, y) = -b(x, y)\sin[\psi(x, y)], \tag{10}$$

and the 2-D complex fringe pattern[55] would be

$$g - i\hat{g} = b \exp(i\psi). \tag{11}$$

If we have a fringe pattern image $g(x, y)$ that has a unique orientation angle $\beta(x, y)$ associated with each point, as shown in Fig. 3(a), then we find that our Fourier spiral phase operator has the following effect on the fringe pattern:

$$\mathcal{F}^{-1}\{\exp[i\phi(u, v)]\mathcal{F}\{b(x, y)\cos[\psi(x, y)]\}\} \cong i \exp[i\beta(x, y)]b(x, y)\sin[\psi(x, y)]. \tag{12}$$

The approximation is exact for the particular case of perfectly straight fringe patterns. This result can be demonstrated by considering a straight, equispaced fringe pattern with constant modulation b_0:

$$g_1(x, y) = b_0 \cos[2\pi(u_0 x + v_0 y)]. \tag{13}$$

The FT of this pattern is a pair of delta functions

$$\mathcal{F}\{g_1(x, y)\} = G_1(u, v) = \frac{b_0}{2}[\delta(u - u_0, v - v_0) + \delta(u + u_0, v + v_0)]. \tag{14}$$

Applying the spiral phase factor and noting that $\tan(\beta_0) = v_0/u_0$, we obtain

$$\exp(i\phi)G_1(u, v) = \exp(i\beta_0)\frac{b_0}{2}[\delta(u - u_0, v - v_0) - \delta(u + u_0, v + v_0)]. \tag{15}$$

Finally, taking the inverse FT gives the exact expression

$$\mathcal{F}^{-1}\{\exp(i\phi)\mathcal{F}\{b_0 \cos[2\pi(u_0 x + v_0 y)]\}\} = i \exp(i\beta_0)b_0 \sin[2\pi(u_0 x + v_0 y)]. \tag{16}$$

The heuristic derivation of the more general approximation (12) is as follows. If we consider the FT of the localized fringe pattern in Fig. 3(a), then we obtain the distribution shown in Fig. 3(b). If we try to localize the pattern to a small region, then the spectral lobes in Fig. 3(b) become larger. The formal mathematical description of this process, utilizing the method of stationary phase, is described in Paper II. We ignore here the uncertainty principle, which limits localizing the fringes and the fringe lobes simultaneously. A cosine fringe pattern will give two lobes with the same polarity. Multiplying the lobes by the spiral phase will change the lobes to opposite polarity and introduce a phase factor $\exp(i\beta)$. It is well-known that changing lobe polarity changes a cosine into a negative sine, and the orientational phase then appears as a factor on top of this quadrature effect.

The above Fourier multiplication is also equivalent to a 2-D convolution of the image function with a spatial spiral phase kernel function $s(x, y)$:

$$\mathcal{F}^{-1}\{\exp[i\phi(u, v)]\mathcal{F}\{g(x, y)\}\} = g(x, y) \;**\; s(x, y), \tag{17}$$

where the kernel function can be shown by general Fourier techniques to be a rather interesting spiral phase, inverse-square function,

$$s(x, y) = \frac{i(x + iy)}{2\pi(x^2 + y^2)^{3/2}} = \frac{i \exp(i\theta)}{2\pi r^2}. \tag{18}$$

The spatial polar coordinates are defined as usual:

$$x = r\cos(\theta), \qquad y = r\sin(\theta). \tag{19}$$

Equation (18) can also be interpreted as the complex sum of two 2-D Riesz kernels[29]:

$$\mathcal{F}^{-1}\left\{-i\frac{u}{q}\right\} = \frac{x}{2\pi r^3}, \qquad \mathcal{F}^{-1}\left\{-i\frac{v}{q}\right\} = \frac{y}{2\pi r^3}, \tag{20}$$

where the spectral radial coordinate q is defined by

$$q^2 = u^2 + v^2. \tag{21}$$

These real functions are more familiar as the singular kernels in singular integral theory, but they can be combined in the real and imaginary parts of the complex Riesz kernel of Eq. (18). The complex notation is unusable in dimensions greater than two but can be advantageous in two dimensions because it implies possible optical implementations. The convolution kernel approach may be important for efficient implementations of the spiral phase transformation. The kernel clearly shows that the spiral phase transform is nonlocal, with a variation $(1/r^2)$ rather like that of the nonlocal 1-D HT kernel $(1/x)$:

$$\mathcal{F}^{-1}\{i\ \mathrm{sign}(u)\} = \frac{1}{\pi x}. \tag{22}$$

We have defined an orientational phase factor $\exp[i\beta(x,y)]$ that is simply related to the fringe angle $\beta(x,y)$. The Fourier spiral phase approximation is derived by considering the Fourier components of localized fringes and is valid for suitably smoothly varying parameters (the local simplicity constraint[39]). Initial experiments indicate that the accuracy is better than 1% for typical patterns, where the fringe radius of curvature is greater than the fringe spacing. In fact, relation (12) can actually be used to define the orientation $\beta(x,y)$, but we take an alternative approach in the following examples. Orientation estimators are of great interest in human and computer vision, with some reliable methods currently available.[39] A number of other researchers have noted the importance of fringe orientation in fringe demodulation.[56,57] We use a special orientation estimator to find β_e, an estimator that does not flip 180° from fringe to fringe (i.e., it is not a simple gradient estimator) and so maintains local continuity. Details of orientation estimation are provided in the appendix to Paper II. The next step is simply to extract $\hat{g}$ and calculate the 2-D complex image:

$$g - i\hat{g} = g - \exp(-i\beta_e)\mathcal{F}^{-1}\{\exp(i\phi)\mathcal{F}\{g\}\}. \tag{23}$$

The process can be seen as a combination of pure phase function multiplication in the space domain (x,y) and in the Fourier domain (u,v). The operator $\mathcal{V}\{\}$, defined by

$$\mathcal{V}\{g\} = -i\exp(-i\beta)\mathcal{F}^{-1}\{\exp(i\phi)\mathcal{F}\{g\}\}, \tag{24}$$

shall be referred to as the vortex operator for brevity in the following text. The vortex operator has the following invariant properties: scale, translation, and rotation (for properly defined β). In essence, the operator satisfies all the requirements of a hypothetical 2-D quadrature transform. The demodulation process defined by $\mathcal{V}\{\}$ can be said to be natural in the sense that $b\sin(\psi)$ [or, more correctly, $-b\sin(\psi)$] is the natural quadrature of $b\cos(\psi)$ for 2-D functions as well as 1-D functions.

It transpires that an accurate orientation estimate is not necessary for high-accuracy phase demodulation or for high-accuracy amplitude demodulation, as can be seen in the following case, where we consider an orientation with an error $\epsilon = \epsilon(x,y)$. The vortex operator then gives

$$\begin{aligned}\mathcal{V}_\epsilon\{g\} &= -i\exp[-i(\beta+\epsilon)]\mathcal{F}^{-1}\{\exp(i\phi)\mathcal{F}\{g\}\}\\ &= b(x,y)\sin[i\psi(x,y)]\exp(-i\epsilon).\end{aligned} \tag{25}$$

For small values of the error ($|\epsilon| < 0.1$),

$$\mathcal{V}_\epsilon\{g\} \approx b(x,y)\sin[i\psi(x,y)]\left(1 - i\epsilon - \frac{\epsilon^2}{2}\right). \tag{26}$$

The phase estimate in this case ignores the imaginary part of the vortex transform:

$$\tan[\psi_\epsilon(x,y)] = \frac{\Re[\mathcal{V}\{g(x,y)\}]}{g(x,y)} = \left(1 - \frac{\epsilon^2}{2}\right)\tan[\psi(x,y)]. \tag{27}$$

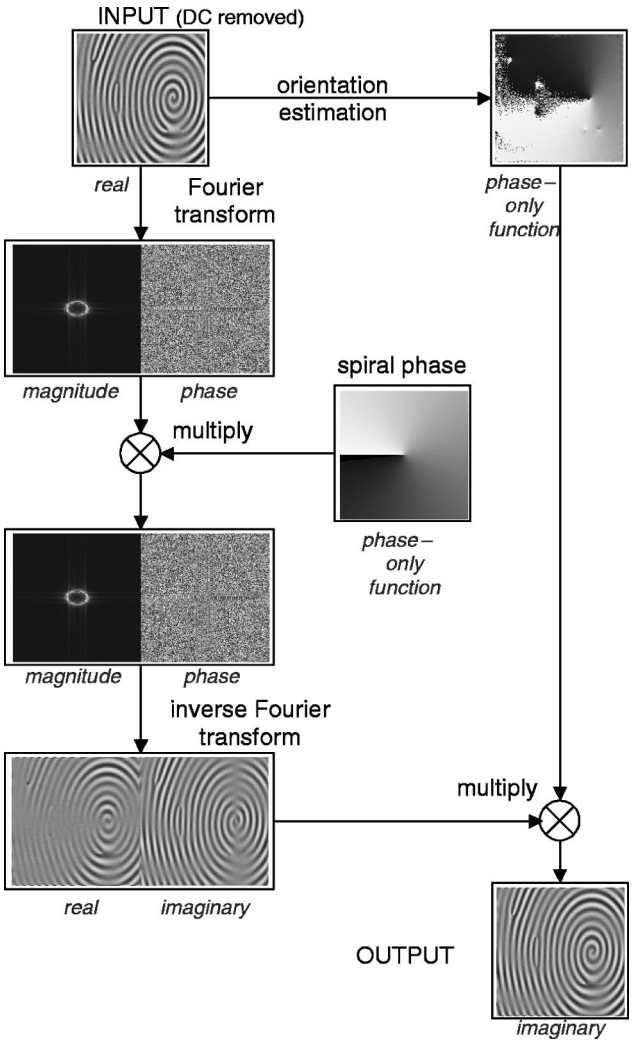

Fig. 4. Spiral phase algorithm for 2-D quadrature function estimation. A purely real fringe pattern is converted to a purely imaginary quadrature pattern by the spiral phase transform. The pattern shown could be an interferogram or a fingerprint pattern, for example. The operation consists of a frequency-domain spiral phase multiplication followed by a space-domain orientation phase multiplication. An additional multiplication can convert the output to real.

The error in the phase estimate, $\delta\psi = \psi_\epsilon - \psi$, then has a particularly simple form[58]:

$$\delta\psi(x,y) \approx -\frac{\epsilon^2}{4}\sin[2\psi(x,y)]. \tag{28}$$

So the phase error is second order and follows the classic second-harmonic pattern familiar in phase-shifting interferometry. If, for example, the orientation error were a rather poor 0.1 rad, then the demodulated phase would be in error by a maximum of 0.0025 rad. The amplitude error derived from the real part of relation (26) is $\delta b = b_\epsilon - b$, where

$${b_\epsilon}^2 \approx b^2\cos^2(\psi) + b^2\left(1 - \frac{\epsilon^2}{2}\right)^2\sin^2(\psi), \tag{29}$$

$$\delta b \approx -\frac{\epsilon^2}{2}b\sin^2(\psi) = -\frac{\epsilon^2}{4}b[1 - \cos(2\psi)]. \tag{30}$$

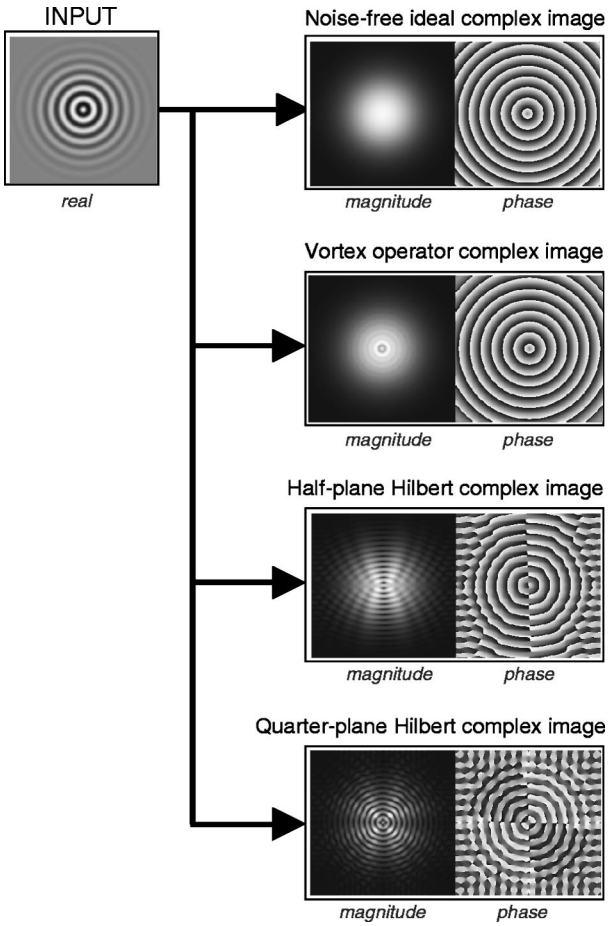

Fig. 5. Comparison of 2-D Hilbert transform (HT) methods for a simple image. The complex images generated by three different methods are compared with the ideal. The vortex operator gives a result similar to the ideal magnitude and phase. Both the half-plane and quarter-plane Hilbert operators give highly anisotropic estimates of magnitude and phase.

The error is again second order with respect to the orientation error with dc and second-harmonic terms. However, by reformulating Eq. (23), we can demodulate the amplitude without any errors that are due to the orientation error:

$$|g|^2 + |\hat{g}|^2 = |g|^2 + |\mathcal{F}^{-1}\{\exp(i\phi)\mathcal{F}\{g\}\}|^2 = |b|^2 \Rightarrow \delta b \equiv 0. \quad (31)$$

In summary, the errors in the orientation estimate cause only second-order errors in the demodulated phase. The demodulated amplitude may have second-order errors or no errors at all, depending on the demodulation algorithm. The second-order error property is indeed fortunate and makes the vortex transform inherently robust to errors.

A particularly elegant solution arises for circular symmetric patterns such as circular fringes. In this case the orientational phase in Eq. (24) is just a single spiral but in the opposite sense of the Fourier spiral—the overall transform is then a double spiral (or double vortex) transform. As far as we know, the direct demodulation of simple closed-curve fringe patterns has not been presented before. We shall present examples in Section 4. In fringe patterns with disjoint closed-curve regions (that is to say, several separated regions such as the number 8, for example), the question of global sign choice for the orientation is ambiguous, and each separate region must have a sign allocated arbitrarily (or based on an additional global constraint). For clarity, our examples will be restricted to simple (nondisjoint) closed-curve patterns.

4. APPLICATION TO TWO-DIMENSIONAL FRINGE PATTERNS

The first example using the vortex operator is a fringe pattern that could quite easily occur in a human fingerprint or as an optical interference pattern. Figure 4 is a flow chart of our proposed vortex operator algorithm. Note that the orientation estimation step produces the pure phase function $\exp(i\beta)$, which is discussed in detail after Eq. (22). The direct algorithm is computationally efficient, requiring just two FTs and two multiplications (and an additional orientation estimation with comparable computational complexity). In the examples shown the offset function $a(x, y)$ is very slowly varying and easily removed, but in general the removal requires more sophisticated processing. The output function is in quadrature to the input. All the examples are computed by using the fast Fourier transform on 128×128 discrete images without any additional windowing operations. Note how the FT magnitude contains a spectral lobe that is a continuous ring. Such rings have made previous attempts at direct FT demodulation impossible.[46,59]

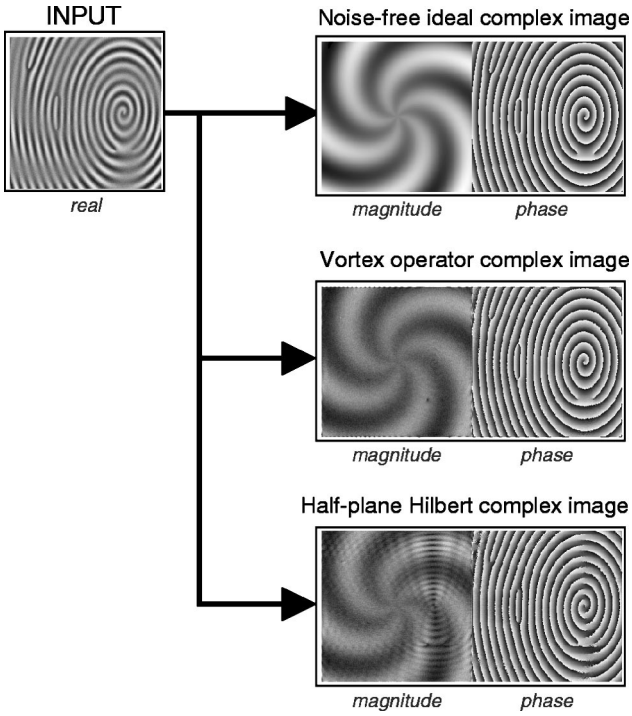

Fig. 6. Comparison of 2-D HT methods for an AM–FM image. The input image has both amplitude and frequency modulation with 10% uniform random noise added. The complex image generated by the vortex operator is visually (and numerically) close to the ideal. In contrast, the half-plane Hilbert result shows gross errors in both magnitude and phase estimates.

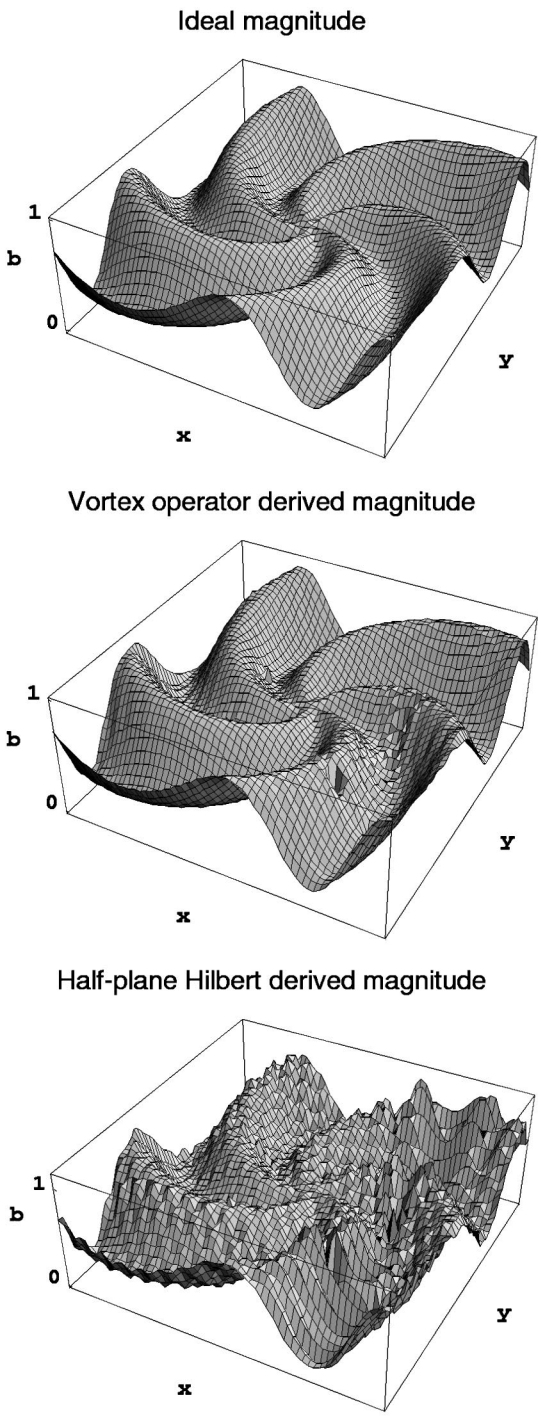

Fig. 7. Comparison of demodulated magnitudes for vortex operator and half-plane Hilbert operator. The vortex operator derives a close estimate of the complex image magnitude, failing only at discontinuities in the phase. The half-plane Hilbert operator derives a highly oscillatory estimate of the magnitude with substantial errors in all regions.

In Fig. 5 we show details of the input and output images compared with idealized quadrature pairs and alternative demodulation methods. The quadrant "Hilbert" transform[3,14] has been included because it appeared recently in a problematic definition of the multidimensional analytic signal.[11] The vortex operator complex image is visually very close to the ideal, failing only close to the discontinuity at the center. The half-plane Hilbert method fails seriously for any closed-curve fringe pattern. The horizontal fringes are highly distorted in this case, leading to a dark region in the estimated magnitude. The phase shows a local sign ambiguity and is also highly distorted in the transition region.[10] The quadrant HT gives a slightly more isotropic estimate of the magnitude but fails rather badly with the phase estimate.[11]

In Fig. 6 we show a comparison of methods applied to the interference pattern from Fig. 4. This time the initial image has both amplitude and phase (AM–FM) structure.[60] The input image also has uniform random noise added for realism (10% of peak signal, 100% of minimum signal). Again, the vortex operator generates an estimated complex image that is visually close to the ideal. The measured relative amplitude error $\delta b/b$ is in the range -28% to $+10\%$ within a region less than half of a fringe from the central discontinuity of the underlying conical phase function. Outside a region just two fringes from the center, the measured relative error drops below 3%, and the closeness to the ideal magnitude is clear in Fig. 7. The errors appear to be related to extremes of fringe curvature and spacing. In contrast, the half-plane HT produces anisotropic magnitude and phase estimates with the usual visible artefacts, as seen in Fig. 6. Note that the half-plane artefacts are typically very large (with measured magnitude errors in the range -98% to $+35\%$, for example) and widely dispersed, as illustrated in Fig. 7.

We believe that the general demodulation methods developed here for fringe patterns may be extended to more general 2-D patterns. However, the local orientation can no longer be defined (local simplicity does not apply), and the full RT approach with its additional components must be utilized. In a similar manner, extensions to three or more dimensions are possible.

5. EXACT SOLUTION FOR CIRCULAR SYMMETRIC PATTERNS

Our equation defining the vortex operator derives from an approximation linking the sine and cosine components of a general fringe pattern. Equation (16) gives an exact solution for equispaced straight fringes, but equispaced circular fringes (as comprehensively analyzed by Amidror[61]) transform only approximately as follows:

$$\mathcal{V}\{b(r)\cos(\lambda r)\} \cong b(r)\sin(\lambda r), \qquad \beta = \theta, \ 0 < \lambda. \tag{32}$$

As mentioned in Section 3, the transform is a particularly simple double vortex transform. However, the vortex transform generally breaks down near the origin, where the phase has a conical discontinuity, as shown in Fig. 5. Nevertheless, there is at least one simple circular symmetric function that transforms exactly with use of the vortex operator:

$$\mathcal{V}\{J_0(\lambda r)\} \equiv J_1(\lambda r), \quad 0 < \lambda. \tag{33}$$

This relation can be derived directly from the 2-D Fourier properties of Bessel functions (given on p. 661 of Bracewell's 2-D imaging book[62]). The Bessel functions asymptotically approach decaying sinusoids for $\lambda r > 3$, which is essentially within one fringe period. The exact solution above suggests an inverse formula:

$$\mathcal{V}^{-1}\{J_1(\lambda r)\} = J_0(\lambda r), \tag{34}$$

where

$$\mathcal{V}^{-1}\{g\} = i\mathcal{F}^{-1}\{\exp(-i\phi)\mathcal{F}\{\exp(i\beta)g\}\}.$$

However, such an inversion may be impractical in more general, noncircular symmetric cases. This is because the inverse requires that orientation estimation and multiplication take place before the Fourier spiral phase transformation. Consequently, any errors in the orientation, especially discontinuities, will spread widely in the final result. The proposed forward algorithm does not have such a problem; any errors in the orientation estimate remain localized.

6. SUMMARY

We have surveyed the literature on multidimensional Hilbert transforms (HTs) and found that a number of groups have independently adopted a Riesz-transform-based definition (without necessarily recognizing the Riesz transform as such). In other areas of research, the directional, orthant-based HT definitions may have inhibited the evolution of isotropic forms. The proposed formalism for the vortex operator allows a quadrature transform and a complex "analytic" signal to be defined uniquely for any 2-D signals, such as fringe patterns, that satisfy the local simplicity constraint. The constraint is in keeping with the restricted definition of instantaneous frequency and the analytic signal in one dimension.[63] We have demonstrated a new form of fringe pattern analysis by using the vortex operator, which directly demodulates 2-D patterns previously considered impossible. Both amplitude and phase demodulations are facilitated by this technique. We expect the vortex operator and the associated 2-D complex signal to have wide applications beyond the remarkable fringe pattern demodulation presented here. The vortex operator may be implemented simply in an optical system by using a spiral phase plate in the Fourier plane or the back focal plane of an imaging system, allowing near-instantaneous evaluation of quadrature functions and suggesting new optical imaging modes.

Note added in proof. Soon after this paper was submitted, Felsberg and Sommer[64] presented a paper that corrected the previously anisotropic results of Bülow and Sommer.[11] They introduce the so-called "monogenic" function as the extension of the analytic function to n dimensions. The idea of a monogenic function was taken from an area known as geometric algebra, closely connected with the Clifford algebra and quaternions. In a paper by Gull *et al.*,[65] the monogenic function is derived from the Cauchy–Riemann conditions in n dimensions by using geometric algebra.

Address correspondence to Kieran G. Larkin at the first address given on the title page or by e-mail, kieran@research.canon.com.au.

REFERENCES AND NOTES

1. D. Gabor, "Theory of communications," J. Inst. Electr. Eng. **93**, 429–457 (1947).
2. S. Lowenthal and Y. Belvaux, "Observation of phase objects by optically processed Hilbert transform," Appl. Phys. Lett. **11**, 49–51 (1967).
3. H. Stark, "An extension to the Hilbert transform product theorem," Proc. IEEE **59**, 1359–1360 (1971).
4. J. K. T. Eu and A. W. Lohmann, "Isotropic Hilbert spatial filtering," Opt. Commun. **9**, 257–262 (1973).
5. J. Ojeda-Castanada and E. Jara, "Isotropic Hilbert transform by anisotropic spatial filtering," Appl. Opt. **25**, 4035–4038 (1986).
6. E. Peli, "Hilbert transform pairs mechanisms," Invest. Ophthalmol. Visual Sci. **30** (ARVO Suppl.), 110 (1989).
7. A. W. Lohmann, E. Tepichin, and J. G. Ramirez, "Optical implementation of the fractional Hilbert transform for two-dimensional objects," Appl. Opt. **36**, 6620–6626 (1997).
8. Y. M. Zhu, F. Peyrin, and R. Goutte, "The use of a two-dimensional Hilbert transform for Wigner analysis of 2-dimensional real signals," Signal Process. **19**, 205–220 (1990).
9. S. L. Hahn, *Hilbert Transforms in Signal Processing* (Artech House, Norwood, Mass., 1996).
10. J. P. Havlicek, J. W. Havlicek, and A. C. Bovik, "The analytic image," in *Proceedings of the IEEE International Conference on Image Processing* (IEEE, New York, 1997), Vol. 2, pp. 446–449.
11. T. Bülow and G. Sommer, "A novel approach to the 2D analytic signal," presented at the 8th International Conference on Computer Analysis of Images and Patterns, Ljubljana, Slovenia, September 1–3, 1999.
12. T. Bülow, "Hypercomplex spectral signal representations for the processing and analysis of images," Ph.D. dissertation (Christian Albrechts University, Kiel, Germany, 1999).
13. M. Craig, "Analytic signals for multivariate data," Math. Geol. **28**, 315–329 (1996).
14. M. A. Fiddy, "The role of analyticity in image recovery," in *Image Recovery: Theory and Application*, H. Stark, ed. (Academic, Orlando, Fla., 1987).
15. J. R. Fienup and C. C. Wackerman, "Phase retrieval stagnation problems and solution," J. Opt. Soc. Am. A **3**, 1897–1907 (1986).
16. Not surprisingly, a separable (i.e., orthant) definition of the multidimensional HT leads only to separable solutions.
17. D. G. Fulton and G. Y. Rainich, "Generalisations to higher dimensions of the Cauchy integral formula," Am. J. Math. **54**, 235–241 (1932).
18. M. R. Arnison, C. J. Cogswell, N. I. Smith, P. W. Fekete, and K. G. Larkin, "Using the Hilbert transform for 3D visualisation of differential interference contrast microscope images," J. Microsc. **199**, 79–84 (2000).
19. K. G. Larkin, "Efficient nonlinear algorithm for envelope detection in white light interferometry," J. Opt. Soc. Am. A **13**, 832–843 (1996).
20. According to Ahmed Zayed [A. I. Zayed, *Handbook of Generalized Function Transformations* (CRC Press, Boca Raton, Fla., 1996)], the HT was so named by G. H. Hardy after David Hilbert (1862–1943), who was the first to observe the conjugate functions now known as a HT pair in D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* (Tuebner, Leipzig, 1912).
21. M. Riesz, "Sur les fonctions conjuguées," Math. Z. **27**, 218–244 (1927).
22. F. G. Tricomi, "Equazioni integrali contenenti il valor principale di un integrale doppio," Math. Z. **27**, 87–133 (1928).
23. G. Giraud, "Sur une classe generale d'equations a integrales principales," C. R. Acad. Sci. **202**, 2124–2126 (1936).

24. S. G. Mikhlin, "Singular integral equations," Usp. Mat. Nauk **3**, 29–112 (1948) (in Russian).
25. S. G. Mikhlin, *Multidimensional Singular Integrals and Integral Equations* (Pergamon, Oxford, UK, 1965).
26. A. P. Calderon and A. Zygmund, "On the existence of certain singular integrals," Acta Math. **88**, 85–139 (1952).
27. A. Zygmund, "On singular integrals," Rend. Mat. **16**, 468–505 (1957).
28. There may have been some confusion about the priority of crucial results in the properties of multidimensional singular integrals. It seems that in 1948 Mikhlin[24,25] showed the L^2 boundedness of the 2-D RT, whereas in 1952 Calderon and Zygmund[26,27] proved the more general L^p boundedness of the n-dimensional RT.
29. E. M. Stein, *Singular Integrals and Differentiability Properties of Functions* (Princeton U. Press, Princeton, N.J., 1970).
30. A. Carbery, "Harmonic analysis of the Calderon–Zygmund school, 1970–1993," Bull. London Math. Soc. **30**, 11–23 (1998).
31. M. N. Nabighian, "The analytic signal of two-dimensional magnetic bodies with polygonal cross-section: its properties and use for automated anomaly interpretation," Geophysics **37**, 507–517 (1972).
32. M. N. Nabighian, "Toward a three-dimensional automatic interpretation of potential field data via generalized Hilbert transform: fundamental relations," Geophysics **49**, 780–786 (1984).
33. W. M. Moon, A. Ushah, V. Singh, and B. Bruce, "Application of 2-D Hilbert transform in geophysical imaging with potential field data," IEEE Trans. Geosci. Remote Sens. **26**, 502–510 (1988).
34. A. E. Barnes, "Theory of 2-D complex seismic trace analysis," Geophysics **61**, 264–272 (1996).
35. C. Pudney and M. Robbins, "Surface extraction from 3D images via local energy and ridge tracing," in *Digital Image Computing: Techniques and Analysis* (Australian Pattern Recognition Society, Brisbane, Australia, 1995), pp. 240–245.
36. M. C. Morrone and D. C. Burr, "Feature detection in human vision: a phase-dependent energy model," Proc. R. Soc. London Ser. B **235**, 221–245 (1988).
37. E. Peli, "Contrast in complex images," J. Opt. Soc. Am. A **7**, 2032–2040 (1990).
38. P. Kube and P. Perona, "Scale-space properties of quadratic feature detectors," IEEE Trans. Pattern Anal. Mach. Intell. **18**, 987–999 (1996).
39. G. H. Granlund and H. Knutsson, *Signal Processing for Computer Vision* (Kluwer Academic, Dordrecht, The Netherlands, 1995).
40. R. Muller and J. Marquard, "Die Hilberttransformation und ihre Verallgemeinerung in Optik und Bildeverarbeitung," Optik (Stuttgart) **110**, 99–109 (1999).
41. J. A. Davis, D. E. McNamara, D. Cottrel, and J. Campos, "Image processing with the radial Hilbert transform: theory and experiments," Opt. Lett. **25**, 99–101 (2000).
42. R. N. Bracewell, *The Fourier Transform and Its Applications* (McGraw-Hill, New York, 1978).
43. M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," J. Opt. Soc. Am. **72**, 156–160 (1982).
44. K. A. Nugent, "Interferogram analysis using an accurate fully automatic algorithm," Appl. Opt. **24**, 3101–3105 (1985).
45. D. J. Bone, H.-A. Bachor, and R. J. Sandeman, "Fringe-pattern analysis using a 2-D Fourier transform," Appl. Opt. **25**, 1653–1660 (1986).
46. T. Kreis, *Holographic Interferometry. Principles and Methods* (Akademie, Berlin, 1996), Vol. 1.
47. J. L. Marroquin, J. E. Figueroa, and M. Servin, "Robust quadrature filters," J. Opt. Soc. Am. A **14**, 779–791 (1997).
48. J. F. Nye and M. V. Berry, "Dislocations in wave trains," Proc. R. Soc. London, Ser. A **336**, 165–190 (1974).
49. W. J. Condell, "Fraunhofer diffraction from a circular annular aperture with helical phase factor," J. Opt. Soc. Am. A **2**, 206–208 (1985).
50. P. Coullet, L. Gil, and F. Rocca, "Optical vortices," Opt. Commun. **73**, 403–408 (1989).
51. I. V. Basistiy, V. Y. Bazhenov, M. S. Soskin, and M. V. Vasnetsov, "Optics of light beams with screw dislocations," Opt. Commun. **103**, 422–428 (1993).
52. D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping* (Wiley, New York, 1998).
53. There are a number of ways to remove the offset. Low-pass filtering is the simplest but often not the best method. In situations with multiple phase-shifted interferograms, the difference between any two frames will have the offset nullified. Adaptive filtering methods can also provide more accurate offset removal. In practice, offset removal may be difficult. The difficulty exists even for 1-D signal demodulation using Hilbert techniques, as shown in detail by N. E. Huang, Z. Shen, S. Long, M. C. Wu *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis," Proc. R. Soc. London Ser. A **454**, 903–995 (1998). We shall not discuss the difficulty further in this initial exposition, but it should be noted that failure to remove the offset signal correctly may introduce significant errors.
54. A. H. Nuttall, "On the quadrature approximation to the Hilbert transform of modulated signals," Proc. IEEE **54**, 1458–1459 (1966).
55. We shall refrain from calling this function the 2-D analytic signal at present because there are several conflicting definitions of analyticity in multiple dimensions. The alternative term "monogenic" does not seem appropriate because the word now has another widespread use in molecular genetics.
56. Q. Yu and K. Andresen, "Fringe-orientation maps and fringe skeleton extraction by the two-dimensional derivative-sign binary-fringe method," Appl. Opt. **33**, 6873–6878 (1994).
57. J. G. Daugman and C. J. Downing, "Demodulation, predictive coding, and spatial vision," J. Opt. Soc. Am. A **12**, 641–660 (1995).
58. K. G. Larkin and B. F. Oreb, "Propagation of errors in different phase-shifting algorithms: a special property of the arctangent function," in *Interferometry: Techniques and Analysis*, G. M. Brown, O. Y. Kwon, M. Kujawinska, and G. T. Reid, eds., Proc. SPIE **1755**, 219–227 (1992).
59. D. Malacara, M. Servin, and Z. Malacara, *Interferogram Analysis for Optical Testing* (Marcel Dekker, New York, 1998).
60. Sharp-eyed readers may have noticed that the fringe pattern used in Figs. 4, 6, and 7 actually contains some spiral discontinuities, which are manifested as ridge endings and bifurcations. The fringe pattern actually satisfies the local simplicity constraint everywhere except at the spiral center points. The robustness of the vortex transform to these discontinuities is significant but is not explored further in our initial exposition of the method.
61. I. Amidror, "Fourier spectrum of radially periodic images," J. Opt. Soc. Am. A **14**, 816–826 (1997).
62. R. N. Bracewell, *Two-Dimensional Imaging* (Prentice-Hall, Englewood Cliffs, N.J., 1995).
63. L. Cohen, P. Loughlin, and D. Vakman, "On an ambiguity in the definition of the amplitude and phase of a signal," Signal Process. **79**, 301–307 (1999).
64. M. Felsberg and G. Sommer, "The multidimensional isotropic generalization of quadrature filters in geometric algebra," in *Algebraic Frames for the Perception–Action Cycle*, AFPAC 2000 (Springer-Verlag, Heidelberg, 2000), pp. 175–185.
65. S. Gull, A. Lasenby, and C. Doran, "Imaginary numbers are not real—the geometric algebra of spacetime," Found. Phys. **23**, 1175–1201 (1993).