symmetry

MDPI

Article

Two-Dimensional Hermite Filters Simplify the Description of High-Order Statistics of Natural Images

Qin Hu [1] and Jonathan D. Victor[2,*]

[1] Microsoft Research, One Microsoft Way, Redmond, WA 98052 USA ; huqinpku@hotmail.com
[2] Feil Family Brain and Mind Research Institute, Weill Cornell Medical College, 1300 York Ave., NewYork, NY 10065 USA; jdvicto@med.cornell.edu
* Correspondence: jdvicto@med.cornell.edu; Tel.: +1-212-746-2343

Academic Editor: name
Received: ?? date; Accepted: date; Published: date

Abstract: Natural image statistics play a crucial role in shaping biological visual systems, understanding their function and design principles, and designing effective computer-vision algorithms. High-order statistics are critical for conveying local features, but they are challenging to study – largely because their number and variety is large. Here, via the use of two-dimensional Hermite (TDH) functions, we identify a covert symmetry in high-order statistics of natural images that simplifies this task. This emerges from the structure of TDH functions, which are an orthogonal set of functions that are organized into a hierarchy of ranks. Specifically, we find that the shape (skewness and kurtosis) of the distribution of filter coefficients depends only on the projection of the function onto a 1-dimensional subspace specific to each rank. The characterization of natural image statistics provided by TDH filter coefficients reflects both their phase and amplitude structure, and we suggest an intuitive interpretation for the special subspace within each rank.

Keywords: image statistics; skewness; kurtosis; orthogonal functions; steerable filters

1. Introduction

Achieving a thorough understanding the statistics of our visual environment is important from both a biological point of view and an engineering point of view. The biological relevance is that the statistics of the natural environment are a strong constraint under which visual systems evolve, develop, and function[1]. The engineering relevance is that a knowledge of image statistics is important for many problems in computer vision [2], including image de-noising, image classification [3-6]), image compression, and texture synthesis [7]. However, understanding image statistics is hampered by the simple fact that the space of image statistics is so large. Here we describe some progress in this direction: a specific filter-based approach that identifies a hidden symmetry, providing a simplified description of high-order natural image statistics, specifically, those of order three and four.

The reason for our focus on high-order statistics is that they carry local visual features, such as lines, corners, and edges [8, 9], but – because of the curse of dimensionality –they are challenging to analyze. In contrast, second-order statistics are concisely captured by the power spectrum, because

it is the Fourier transform of the autocorrelation function. As is well-known, the power spectrum of natural images is approximately k^{-2} (where k is spatial frequency) [10, 11]. However, while the power spectrum captures important spatial regularities of natural images – such as distance-independent scaling [12], it is far from a complete statistical description of natural images. For example, a synthetic image consisting of Gaussian noise with a k^{-2} power spectrum looks drastically different from a real natural image, even though the spectra are similar. Conversely, modifying a natural image by flattening its power spectrum but preserving its phases leaves its salient spatial features readily recognizable. Thus, most of the features that make an image look "natural," such as edges and contours, are coded in its phases as well as its Fourier amplitudes [8, 9, 13]. Translated into the spatial domain, these phase correlations correspond to image statistics that are ignored by the power spectrum: joint distributions of image intensities at three or more points, and aspects of the pairwise intensity distributions beyond their variances and covariances.

Since a direct tabulation of the joint distribution of multiple pixel values is impractical, a natural strategy is to focus on specific univariate distributions – namely, the distribution of outputs of filters ("filter coefficients") placed on images. Typically, this approach is implemented with filter profiles that have a prominent orientation and dominant spatial frequency – either Gabor functions or Gabor-like wavelets, a choice motivated by concepts of visual processing and independent components analysis of natural images[14, 15]. For natural images, the distributions of wavelet coefficients are highly kurtotic, having sharp peaks and much longer tails compared to a Gaussian distribution with the same variance [16]. Interestingly, [3] showed that this could be used to distinguish natural images from synthetic ones (including realistic computer-generated scenes), by applying linear classifiers to a feature space of wavelet coefficients. Other investigators have also used wavelet coefficients as a starting point, but focused on the extent to which wavelet coefficients are independent [17, 18]. Thus, the filter approach provides a useful characterization of natural image statistics -- but even with a filter-based approach, the number of parameters required to describe high-order image statistics is still large: a two-dimensional basis set is a two-parameter family.

Here we show that the description of these filter coefficient distributions is simplified when, instead of Gabor-like filters, we use the two-dimensional Hermite functions (TDH's) as filters. TDH functions [19-25] form an orthonormal basis that is halfway between the pixel basis and the Fourier basis, and their shapes are quite different from that of Gabor-like filters or one-dimensional wavelets.

We note that symmetry plays two distinct roles in this study: the purely mathematical symmetry properties of the TDH functions, and the empirical finding that they reveal a hidden symmetry in the statistics of natural images. Specifically, the TDH functions at each rank form a representation of the surfaces of spheres of progressively ascending dimensions: the functions of rank 2 correspond to the points on the surface of an ordinary sphere; the functions of rank 3 correspond to the points on the surface of a hypersphere, etc. The statistics of their filter coefficients – in particular, their skewnesses and kurtoses --may therefore be regarded as functions on these spheres. *A priori*, these functions could have any behavior, but we find that their behavior is surprisingly simple: they are either constant, or depend only on the projection onto a single axis. This simplification depends on both the phase and amplitude characteristics of natural scenes, and, critically, encompasses the distribution of filter coefficients of nonstandard combinations of TDH functions (see Figure 7) that do not have Cartesian or polar symmetry.

2. Materials and Methods

2.1. Two-dimensional Hermite functions: definition and properties

We analyze image statistics via the distribution of values that result from filtering them with two-dimensional Hermite (TDH) functions. Symmetry thus plays two roles in this work: first, the

intrinsic symmetries of the TDH functions themselves, and second, an empirical symmetry of natural image statistics that emerges from this analysis.

We first describe the mathematical properties of TDH functions, with a focus on their symmetries. TDH's (Figure 1) are a set of two-dimensional functions consisting of a product of Hermite polynomials multiplied by a Gaussian envelope. Like wavelets, they are filter functions that are limited in space and spatial frequency. However, they have several other mathematical properties, including additional symmetries. First, the TDH's are symmetrical with respect to space and spatial frequency: other than a multiplicative constant, each TDH is its own Fourier transform. Second, they are orthonormal functions, and as a set, form a complete basis set for functions of two variables. Third, the TDH's are grouped into "ranks": the sole member of the zeroth rank is an ordinary Gaussian; higher rank ranks contain functions of increasing spatial complexity. Finally, within each rank, the TDH's have an extended steerability property. This includes ordinary steerability – the filters can be rotated by forming simple linear combinations – but also, linear combinations within rank provide equivalent basis sets that are separable in Cartesian coordinates (see rows of Figure 1).

Below we define these functions in abstract terms and then give an explicit expression for their polynomial portions; the former makes their key properties transparent, while the latter is necessary for computation. For further details on this approach, see [25]; other descriptions of the properties of these functions in the context of image processing may be found in [19-24].

Taking inspiration from [26, 27], we define the TDH's as the eigenvectors of the operator $D^{1/2}BD^{1/2}$, where D consists of spatial windowing by a two-dimensional Gaussian function (i.e., pointwise multiplication), and B consists of filtering by a two-dimensional Gaussian spatial frequency window (i.e., pointwise multiplication in the spatial frequency domain). Note that D is diagonal in the natural (pointwise spatial) basis, since it consists of pointwise multiplication by the Gaussian; similarly, B is diagonal in the Fourier basis, since it consists of pointwise multiplication by a Gaussian function of spatial frequency. Since the multiplying factors in both cases are positive real numbers, both operators have a naturally-defined principal square root, which we denote as $D^{1/2}$ and $B^{1/2}$. Based on these and other considerations, it can be shown that the operator $D^{1/2}BD^{1/2}$ is self-adjoint, and has a discrete set of eigenvalues [25]. The approach of [28] shows that the eigenvalues are of the form $\lambda = \eta^{1+r}$, for a positive constant $\eta < 1$, where the rank, r, ranges over the non-negative integers [25]. It also shows that the r th rank contains $r+1$ linearly independent functions [25]. Note that this setup is symmetric under interchange of space and spatial frequency, i.e., under interchange of D and B, so the above properties (and those mentioned below) also hold for $B^{1/2}DB^{1/2}$.

Since D corresponds to confinement in space, and B corresponds to confinement in spatial frequency, a TDH function f has the property that successive windowing in space and spatial frequency results in multiplication by a constant (the eigenvalue λ): $D^{1/2}BD^{1/2}f = \lambda f$. That is, for functions f corresponding to eigenvalues λ close to 1, these windowing operations have a small effect – which formalizes the notion that f is confined in both space and spatial frequency.

The eigenfunction of largest eigenvalue (i.e., the TDH function of rank $r = 0$) is a Gaussian, and its eigenvalue is given by $\eta = \left(\frac{2c}{1+\sqrt{1+4c^2}} \right)^2$, where c is the product of the standard deviation of the Gaussians that define the projections of D or B on either coordinate axis.

Since the eigenvalues are all of the form $\lambda = \eta^{1+r}$, the TDH function of rank $r = 0$ has the eigenvalue that is closest to 1, and is therefore the most confined. Successive ranks have exponentially-declining eigenvalues, and are therefore progressively less confined (i.e., is more extensive spatially and contains a progressively broader range of spatial frequencies). TDH functions at different ranks are orthogonal, since they correspond to different eigenvalues of the self-adjoint operator $D^{1/2}BD^{1/2}$.

The extended steerability of the TDH functions is a consequence of combining this setup with the fact that a circularly-symmetric Gaussian is separable both in Cartesian and polar coordinates. As a consequence, both D and B have polar symmetry and separability in Cartesian coordinates, These symmetries are inherited by $D^{1/2}BD^{1/2}$ as well, and must be retained by the eigenspaces, so the existence of Cartesian and polar-symmetric eigenvectors are guaranteed. Since any set of $r+1$ linearly independent eigenvectors forms a basis for each rank, it follows that we can express the Cartesian and polar basis sets as linear combinations of each other.

The second role played by symmetry – the empirical symmetry identified in natural images – is distinct from the spatial symmetries of the Cartesian or polar TDH functions themselves. Rather, this it emerges from an analysis that is motivated by the eigenstructure of the operator $D^{1/2}BD^{1/2}$ (or $B^{1/2}DB^{1/2}$). Since the eigenspace of rank r has dimension $r+1$, the complete set of unit-norm eigenvectors for each eigenvalue can be considered as points on an *r*-sphere (i.e., the surface of an ordinary sphere for $r=2$, or of a hypersphere for $r=3$, etc.) This spherical surface includes not only the Cartesian and polar basis sets, but other eigenfunctions (see Figure 7) that are mixtures of the two, and have no intrinsic symmetry. Descriptors of filter functions distributions (such as the skewness and kurotosis) for the complete set of eigenvectors can thus be viewed as functions on these spheres. While these functions could have any behavior, we find instead that they depend only on the projection onto a single axis – even for filter shapes that are highly irregular.

2.2. Two-dimensional Hermite functions: explicit expressions

As described in §2.1, there are two natural basis sets for the TDH functions of rank r: polar and Cartesian. The polar basis functions are specified by their rotational symmetry (an integer μ, for which a rotation by $2\pi/\mu$ leaves the function unchanged) and the number of zero-crossings along each radius (an integer ν). These indices are related to the rank r by $r=\mu+2\nu$. For $\mu>0$, the basis functions form "cosine" and "sine" pairs:

$$A^{\cos}_{\mu,\nu,\sigma}(R,\theta)=\frac{K}{\sigma}\cos(\mu\theta)\left(\frac{R}{\sigma}\right)^{\mu}P_{\mu,\nu}\left(\frac{R^2}{\sigma^2}\right)\exp\left(-\frac{R^2}{4\sigma^2}\right) \tag{1}$$

and

$$A^{\sin}_{\mu,\nu,\sigma}(R,\theta)=\frac{K}{\sigma}\sin(\mu\theta)\left(\frac{R}{\sigma}\right)^{\mu}P_{\mu,\nu}\left(\frac{R^2}{\sigma^2}\right)\exp\left(-\frac{R^2}{4\sigma^2}\right), \tag{2}$$

where σ sets the overall size of the filter set, K is a normalization constant, and $P_{\mu,\nu}(u)$ is a radial polynomial defined by

$$P_{\mu,\nu}(u)=\sum_{p=0}^{\nu}(-2)^{\nu-p}\frac{(\mu+\nu)!\,\nu!}{(\mu+p)!\,p!\,(\nu-p)!}u^{p}\,. \tag{3}$$

For each even ranks, there is also an unpaired basis function, corresponding to $\mu=0$ and $\nu=r/2$. These basis functions have no angular dependence (central column of Figure 1A), and are given by $A^{\cos}_{0,r/2,\sigma}(R,\theta)$.

A typical Cartesian basis function has the appearance of vignetted $(j+1)\times(k+1)$ checkerboard, where there are j vertical zero-crossings, k horizontal crossings, and these indices are related to the rank by $r=j+k$. It is given by

$$C_{j,k,\sigma}(x,y)=\frac{K}{\sigma}h_j\left(\frac{x}{\sigma}\right)h_k\left(\frac{y}{\sigma}\right)\exp\left(-\frac{x^2+y^2}{4\sigma^2}\right)$$

where $h_j(u)$ and $h_k(u)$ are Hermite polynomials, normalized so that they have the generating function

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} h_n(u) = \exp(uz - \frac{z^2}{2}).$$

As detailed in §2.4, we calculate image statistics of natural images filtered by the polar TDH's, and then use steerability to calculate the statistics of images filtered by other TDH's of a given rank, including the Cartesian TDH filters (as indicated in Figure 2) and intermediate ones. Note that this "steerability" is much more than geometric rotation, as it allows for filters of different shapes, including asymmetric ones (see Figures 4 and 7 below) to be represented in terms of a small basis set.

2.3. Natural images

All 4167 images from the van Hateren natural image database [15] (van Hateren & van der Schaaf, 1998) were chosen for analysis. Each image is 1536 by 1024 pixels, with each pixel's intensity represented by a 16-bit unsigned integer, reflecting an effective bit depth of 12. The images mainly contain landscapes and plants, but occasionally manmade objects such as houses appear.

2.4. Analysis

To characterize high-order statistics of natural images, we calculated the skewness and kurtosis (as "excess kurtosis") of the distribution of filter coefficients, i.e., the distribution of values that result from convolving the images with TDH functions. To focus on the structure of the individual scenes (rather than the overall differences across scenes), skewness and kurtosis were calculated individually for each image, and values were then averaged across the image database.

As shown in Figure 3, this calculation was carried out across 7 spatial scales, spaced in approximately octave steps. The smallest scale used was $\sigma = 7/12$ (0.58) pixels and the largest, $\sigma = 511/12$ (42.6) pixels. At each scale, the image was convolved with polar TDH functions of ranks 0-7 (36 filters in all), and the convolution was sampled at points placed in a rectangular grid on the filtered image. Filters centers were separated by 10 pixels for scales 1-5 and 50 pixels for scales 6 and 7. We then calculated the pure and mixed moments of these distributions up to order 4, and used the extended steerability property (detailed below) to go from the moments for the polar TDH functions to the moments for arbitrary TDH functions. From these moments, skewness and kurtosis were then calculated in the standard fashion.

In detail, computation of the skewness and kurtosis for all TDH functions F of rank r were carried out in parallel, as follows. For each image I, we calculated the pure moments for each polar basis function f

$$M_m(f) = \left\langle \left((f * I)(x,y) \right)^m \right\rangle_{x,y} \quad (4)$$

up to $m = 4$, along with the mixed moments for each pair of functions f and f'

$$M_{m,m'}(f,f') = \left\langle \left((f * I)(x,y) \right)^m \left((f' * I)(x,y) \right)^{m'} \right\rangle_{x,y} \quad (5)$$

up to $m + m' = 4$, and, analogously, the mixed moments $M_{1,1,1}(f,f',f'')$, $M_{2,1,1}(f,f',f'')$, and $M_{1,1,1,1}(f,f',f'',f''')$.

To use the extended steerability property, we wrote the filter function F as a linear combination of the polar basis functions of that rank:

$$F(x,y) = \sum_{n=1}^{r+1} b_n f_n(x,y).$$

Therefore, the convolution of F with an image I can be calculated as a linear combination of the convolutions of the basis functions with the image,

$$\left(F * I\right)(x,y) = \sum_{n=1}^{r+1} b_n \left(f_n * I\right)(x,y). \quad (6)$$

Expressions relating the moments of the distribution of the filter coefficients for F to the moments for the basis functions f_n follow via multinomial expansion of (6), using (4)and (5):

$$M_1(F) = \left\langle \left(F * I\right)(x,y) \right\rangle_{x,y} = \sum_n b_n M_1(f_n), \quad (7)$$

$$M_2(F) = \left\langle \left(\left(F * I\right)(x,y)\right)^2 \right\rangle_{x,y} = \sum_n {b_n}^2 M_2(f_n) + 2\sum_{n_1<n_2} b_{n_1} b_{n_2} M_{1,1}(f_{n_1}, f_{n_2}), \quad (8)$$

$$\begin{aligned} M_3(F) = \left\langle \left(\left(F * I\right)(x,y)\right)^3 \right\rangle_{x,y} &= \sum_n {b_n}^3 M_3(f_n) + 3\sum_{n_1\neq n_2} {b_{n_1}}^2 b_{n_2} M_{2,1}(f_{n_1}, f_{n_2}) \\ &+ 6\sum_{n_1<n_2<n_3} b_{n_1} b_{n_2} b_{n_3} M_{1,1,1}(f_{n_1}, f_{n_2}, f_{n_3}) \end{aligned}, \quad (9)$$

and

$$\begin{aligned} M_4(F) = \left\langle \left(\left(F * I\right)(x,y)\right)^4 \right\rangle_{x,y} &= \sum_n {b_n}^4 M_4(f_n) + 4\sum_{n_1\neq n_2} {b_{n_1}}^3 b_{n_2} M_{3,1}(f_{n_1}, f_{n_2}) \\ &+ 6\sum_{n_1<n_2} {b_{n_1}}^2 {b_{n_2}}^2 M_{2,2}(f_{n_1}, f_{n_2}) + 12\sum_{n_1\neq n_2, n_1\neq n_3, n_2<n_3} {b_{n_1}}^2 b_{n_2} b_{n_3} M_{2,1,1}(f_{n_1}, f_{n_2}, f_{n_3}) \\ &+ 24\sum_{n_1<n_2<n_3<n_4} b_{n_1} b_{n_2} b_{n_3} b_{n_4} M_{1,1,1,1}(f_{n_1}, f_{n_2}, f_{n_3}, f_{n_4}) \end{aligned}. \quad (10)$$

As is standard, the cumulants of the distribution of the filter outputs of F are determined from its moments by

$$\kappa_2 = M_2(F) - \left(M_1(F)\right)^2, \quad (11)$$

$$\kappa_3 = 2\left(M_1(F)\right)^3 - 3M_1(F)M_2(F) + M_3(F), \quad (12)$$

and

$$\kappa_4 = -6\left(M_1(F)\right)^4 + 12\left(M_1(F)\right)^2 M_2(F) - 3\left(M_2(F)\right)^2 - 4M_1(F)M_3(F) + M_4(F). \quad (13)$$

Skewness and (excess) kurtosis are ratios of the cumulants:

$$\gamma_3 = \kappa_3 / {\kappa_2}^{3/2}, \quad (14)$$

and

$$\gamma_4 = \kappa_4 / {\kappa_2}^2. \quad (15)$$

3. Results

We characterized the high-order statistics of natural images via the distribution of filter coefficients for two-dimensional Hermite (TDH) functions. We present the findings for rank 2 first because this low rank allows for a detailed visualization, and then turn to higher ranks.

3.1. Statistics of rank-2 TDH filter coefficients for natural images

To visualize the results for rank 2, we note that the full set of rank-2 filters can be regarded as points on the surface of an ordinary sphere (Figure 4). This follows from the general observation that the r th rank of TDH functions is spanned by contains $r+1$ orthonormal filters, so the full set of unit-magnitude filters of rank r (i.e., the full set of unit-magnitude linear combinations of these

$r+1$ basis elements) may be regarded as the surface of a sphere in $(r+1)$-space. In this spherical representation of rank-2 TDH functions shown in Figure 4, the polar filters correspond to one set of orthogonal directions, the Cartesian filters to a second orthogonal set of directions, and intermediate directions correspond to mixtures of polar or Cartesian filters. The latitude (altitude) indicates the size of the projection onto the target-like TDH function. For TDH functions at the same latitude, the azimuth on the sphere corresponds to the orientation (i.e., the in-plane rotation angle) of the filter function.

Figure 5 shows skewness and kurtosis of the distributions for all TDH filters of rank 2, plotted on the filter space shown in Figure 4. Skewness and kurtosis depend strongly on latitude, but are largely independent of orientation, although there is a small dependence of kurtosis at orientation at the two largest scales. Skewness is maximal for the circularly-symmetric (target-like) filters at the poles and is zero for filters on the equator, while kurtosis is minimal for the target-like filters, and is maximal for filters on the equator.

3.2. Statistics of higher-rank TDH filter coefficients for natural images

For higher ranks, a similar visualization strategy is not possible, so we begin with the skewness and kurtosis for each of the filters in the polar basis set (Figure 6). We focus on filter scale 4, the middle of the range studied; other filter scales gave a similar pattern of results.

With regard to skewness (Figure 6A, second column), there is a single polar filter for which skewness is large; for the others, it is close to zero. For even ranks (consistent with the rank-2 results shown in Figure 5), the single polar filter that has a large skewness is the target-like filter $A^{\cos}_{0,r/2}$; this is the only polar filter with a nonzero mean. For odd ranks, the filter with the largest skewness is the filter with a single horizontal inversion axis, $A^{\sin}_{1,(r-1)/2}$; this filter is specifically sensitive to vertical gradients.

With regard to kurtosis (Figure 6A, third column), the pattern is also a simple one. For even ranks (also consistent with Figure 2), kurtosis is uniform for all filters except the target-like one $A^{\cos}_{0,r/2}$, shown as the middle bar of each histogram in the right column); for the target-like filter, kurtosis is approximately half the size of the others. For odd ranks, the kurtosis is large but uniform across all filters. Thus, we find that for each rank, skewness and kurtosis are either uniform across all polar basis functions, or uniform for all basis functions except for one special filter – the odd-rank filter with a single horizontal inversion axis, or the even-rank filter that is target-like.

For completeness, the first column of Figure 6A shows the variance of each filter's outputs. This is large for target-like filters (center filter in even ranks), and small for all other filters, with sine and cosine pairs resulting in similar variances. As variance is a second-order statistic, this behavior is a consequence of the k^{-2} power spectrum of the images.

The simple behavior of skewness and kurtosis for the TDH functions is not merely a consequence of their polar symmetry. To see this, we repeated the analysis of Figure 6A, but with the polar TDH functions replaced by binarized variants, in which positive values of the polynomial component (eq. (3)) are replaced by +1, and negative values by -1. The binarized variants have the same polar symmetry and sine/cosine pairing as the original TDH functions, and, within ranks, are mutually orthogonal as well. However, when the polynomial portions of the TDH functions are replaced by ±1, neither skewness nor kurtosis have the same simple behavior seen in Figure 6A. Specifically, while the skewness and kurtosis vary over a wide range (approximately 0 to 2 for skewness, 10 to 20 for kurtosis) and this substantial variation is captured in a single basis function at each rank for the original TDH functions, it is spread across many basis functions for the modified ones (Figure 6B).

While Figure 6A suggests that skewness and kurtosis of a general TDH filter depends only on its projection onto the special axis, it only examines filters that are orthogonal to the special axis. For oblique directions, it is possible that this result will not hold. The reason that more complex

behavior may arise in oblique directions is that for moments of order 3 and higher, the steering equations (eqs. (9) and (10) in §2.4) include contributions from mixed moments of the polar TDH's.

Figure 7 shows that despite this potential complication, skewness and kurtosis of a TDH filter's output depends chiefly on the projection of the filter onto the single special axis identified in Figure 6A. It is noteworthy that this holds not only for the Cartesian TDH's, but also for generic TDH's – which typically lack rotational symmetry. Moreover, for ranks $r \geq 3$, TDH functions that share the same projection onto this axis are intrinsically different in shape, and are not merely physical rotations of one another.

In sum, within each rank, skewness and kurtosis of the filter coefficient distribution is either uniform, or uniform in all but one direction in filter space. This axis has a simple interpretation – it is either the target-like TDH function, or the single TDH function that is sensitive to a top-to-bottom gradient. In other words, although TDH filter space has a high dimensionality (equal to the rank+1), the behavior of skewness and kurtosis is always low-dimensional – either uniform, or rotationally symmetric. This simplification constitutes a symmetry of natural image statistics, and goes beyond the overt spatial symmetries of the TDH's themselves: on the one hand, it applies to filter functions that lack either Cartesian or polar symmetry (Figure 7); on the other hand, this simplification fails when the polynomial portion of a TDH filter is replaced by ±1 (Figure 6B), even though this replacement retains all of the spatial symmetries of the filters.

Figure 8 uses this finding to describe the distribution of TDH filter coefficients across all spatial scales in a concise manner. Skewness is characterized by its value for the target-like filter at even ranks ($\gamma_{3,target}$, Figure 8A) and for the filter with a single horizontal inversion axis at odd ranks ($\gamma_{3,horiz}$, Figure 8B). $\gamma_{3,target}$ is a decreasing function of scale and rank and $\gamma_{3,horiz}$ is an increasing function of scale, and (except for rank 1) nearly independent of rank. Kurtosis is characterized by its value for the target-like filter at even ranks ($\gamma_{4,target}$, Figure 8C) and by its value for the remaining filters, at both even and odd ranks ($\gamma_{4,non-target}$, Figure 8D). Both kurtosis quantities are decreasing functions of scale and rank. It would be of interest to characterize the scaling behaviors of the skewness and kurtosis parameters more precisely, but this is beyond the scope of the present study.

3.3. Statistics TDH filter coefficients for altered images

To understand the attributes of natural images that underlie the above findings, we carried out parallel analyses for natural images that were manipulated in several ways prior to the determination of filter coefficients.

First, we examined the role of local mean luminance. To do this, we repeated the analysis of Figure 6, but with subtraction of the local mean luminance over a disk of radius 6σ prior to computing TDH filter outputs (Figure 9A). This manipulation eliminated the difference between the kurtosis for the target-like filter and the others, so that kurtosis was uniform within each rank. Subtraction of the local mean reduced, but did not eliminate, the value of the skewness for the target-like filter. As expected, subtraction of the local mean did not change the distributions for the polar TDH filters that were not target-like, since for $\mu \neq 0$, the trigonometric terms in eqs. (1) and (2) necessarily integrate to 0.

To distinguish the roles of spatial frequency content and phase correlations, we analyzed the distribution of filter coefficients for phase-scrambled images and for images that are spectrally flattened. To isolate the role of spatial frequency content, we created phase-scrambled images by randomizing the phases of the Fourier components in the original images. This effectively results in samples of a spatial Gaussian noise whose power spectrum matches that of the original image. As expected, analysis of these images yielded distributions of TDH filter outputs whose variances matched those of the original images, but for which skewness and kurtosis were zero (not shown). This confirms that spatial frequency content alone does not carry the high-order statistics observed in natural images [8].

To isolate the role of phase correlations, we set the Fourier component amplitudes in the original images to unity, but retained their phases. As in Figure 9A, calculation of filter outputs was carried out with subtraction of the local mean, to retain the isotropy of the kurtosis. Other than for the rank-0 filter, this eliminated the skewness (Figure 9B). The kurtosis remains isotropic. Thus, the heavy-tailed nature of the coefficient distributions depends not only on phase, but also on amplitude.

Finally, to determine the role of the luminance distribution, we calculated the filter coefficient distributions for images subjected to manipulation of the pixel histogram: logarithmic transformation, histogram equalization, and transformation of the intensity histogram to a Gaussian, truncated to 2.56 s.d. (Figure 10). All of these reduced both skewness (by approximately a factor of 10) and kurtosis (by approximately a factor of 5), with near-complete elimination of skewness following the logarithmic transformation. Skewness was concentrated in the filter with a single horizontal inversion axis at odd ranks, and kurtosis was approximately constant within rank.

4. Discussion

Here we show that two-dimensional Hermite (TDH) filters, an orthogonal basis set with a high degree of symmetry, simplify the description of high-order statistics of natural images, both locally and over wide areas. The significance of this result is that high-order statistics carry the local features that distinguish natural images from Gaussian processes [3, 8, 17, 18, 29], but they are challenging to analyze because of their high dimensionality. By identifying a hidden symmetry in high-order statistics, TDH functions provide a kind of dimensional reduction, and therefore, a needed simplification. We emphasize that our goal is focused on understanding natural images, not neural computations *per se*. Specifically, we do not intend to suggest that the visual system uses TDH filters; rather, our point is that they simplify the description of the stimulus set that the visual system must grapple with. This application of TDH functions to characterize natural image statistics is distinct from two other applications of them to vision: a body of work in image processing [19, 21, 23, 24] that uses them to extract local features, and neurophysiologic studies that use them as visual stimuli to analyze the properties of neuronal receptive fields[30, 31].

It is worth noting that the TDH filters constitute a set of functions with an unusually high degree of symmetry. They can be written as a product of functions in either Cartesian or polar coordinates, and thus have both rotational symmetry and steerability. The steerability includes not only the ordinary rotational transformations of the plane, but also rotations in the hyperspheres that correspond to each rank of TDH filters. Moreover, other than a constant factor, each TDH filter is its own Fourier transform – an explicit symmetry relating space and spatial frequency.

Our findings can be viewed as building on [17] and [32], which also focus on the high-order image statistics of natural images. Specifically, these authors examined the distributions of outputs of filters acting on whitened natural images, and the joint distributions of outputs of pairs of filters identified by independent components analysis. [17] showed that the joint distribution is approximately circular, and [32] showed that an improved characterization of the joint distribution could be obtained using an L^p-norm, rather than the Euclidean norm. This near-circularity implies that for any filter, the distribution of outputs has a qualitatively similar heavy-tailed shape. The observation that bandpass filter outputs have similar kurtoses has also been made in other studies [33, 34]. However, this is similarity is only a loose approximation: when analyzed quantitatively (e.g., Figure 5 of [17]), the kurtosis of these distributions varied by at least as factor of two. Here, we show that analysis in terms of TDH filters concisely summarizes this variation: at each rank, the kurtosis of a filter's output is determined by its projection onto a specific direction in filter space.

Examination of the polar TDH filters (Figure 1) suggests the reasons that specific axes are singled out. For the even-rank filters, the special axis is the only filter whose mean is nonzero; all other filters necessarily have a mean of zero because of their sinusoidal dependence on angle. Thus, these filters are the ones that are sensitive to the distribution of local luminances, which are well-known to be heavy-tailed in natural images, both in terms of skewness [35, 36] and kurtosis

[37]. For the odd-rank filters, the identified axis has a horizontal mirror-inversion, with large lobes above and below the horizon. Thus, these filters are likely to be highly sensitive to vertical gradients, and thus, the distributions of their outputs will be skewed by the tendency of illumination to come from above. Consistent with these hypotheses, removal of the local mean (Figure 9A) eliminated the distinctive behavior of target-like filter for kurtosis, and reduced its skewness. When the low spatial frequencies were reduced by spectral flattening, the skewness was eliminated for the odd-rank filters as well. Figure 10 provides further evidence that the distinctive kurtosis for the target-like filters is primarily a consequence of luminance distributions, as it is reduced by attenuating the tails of the luminance distribution via log transformation, histogram-equalization, or Gaussianization.

However, the simplification we observe is not simply a consequence of the arrangement of the positive and negative lobes of the TDH filters, and thus, has deeper roots than the overt spatial symmetries of the TDH filters. The evidence for this is that replacing the Hermite polynomial values by ±1, which preserves the arrangement of their lobes, does not result in a similar simplification of the skewness and kurtosis (Figure 6B). Thus, the crucial factor in our findings is the interaction between the polynomial gradations of the TDH's and the properties of natural images.

5. Conclusions

Two-dimensional Hermite filters provide a simple description of third- and fourth-order statistics of natural images across a range of scales. This simplification is a consequence of the high degree of symmetry of this orthogonal basis set, and the phase, amplitude, and luminance characteristics of natural images.

Acknowledgments: We thank Eyal Nitzany and Matthias Bethge for comments on an earlier version of this manuscript. Supported in part by NIH EY07977 and NIH EY09314 to J.D.V.

Author Contributions: J.D.V. and Q.H. designed the experiments; Q.H. carried out the analysis; J.D.V. and Q.H. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

TDH: Two-dimensional Hermite

References

1. Elder, J.H., J. Victor, and S.W. Zucker, *Understanding the statistics of the natural environment and their implications for vision.* Vision Res, 2016. **120**: p. 1-4.
2. Pouli, T., D.W. Cunningham, and E. Reinhard, *Image statistics and their applications in computer graphics.* Proceedings of Eurographics 2010- State of the Art Reports, 2010: p. 83-112.
3. Lyu, S. and H. Farid, *Higher-order wavelet statistics and their application to digital forensics.* IEEE Workshop on Statistical Analysis in Computer Vision, 2003: p. 94-101.
4. Lyu, S.W. and H. Farid, *Steganalysis using higher-order image statistics.* Ieee Transactions on Information Forensics and Security, 2006. **1**(1): p. 111-119.
5. Lyu, S. and H. Farid, *Detecting hidden messages using higher-order statistics and support vector machines.* Information Hiding, 2003. **2578**: p. 340-354.

6. Lyu, S., D. Rockmore, and H. Farid, *A digital technique for art authentication.* Proc Natl Acad Sci U S A, 2004. **101**(49): p. 17006-10.
7. Chainais, P., *Infinitely divisible cascades to model the statistics of natural images.* IEEE Trans Pattern Anal Mach Intell, 2007. **29**(12): p. 2105-19.
8. Oppenheim, A.V. and J.S. Lim, *The importance of phase in signals.* Proc. IEEE, 1981. **69**: p. 529-541.
9. Morrone, M.C. and D.C. Burr, *Feature detection in human vision: a phase-dependent energy model.* Proc R Soc Lond B Biol Sci, 1988. **235**(1280): p. 221-45.
10. Field, D.J., *Relations between the statistics of natural images and the response properties of cortical cells.* J Opt Soc Am [A], 1987. **4**(12): p. 2379-94.
11. Tolhurst, D.J., Y. Tadmor, and T. Chao, *Amplitude spectra of natural images.* Ophthalmic Physiol Opt, 1992. **12**(2): p. 229-32.
12. Ruderman, D.L., *Origins of scaling in natural images.* Vision Res, 1997. **37**(23): p. 3385-98.
13. Tadmor, Y. and D.J. Tolhurst, *Both the phase and the amplitude spectrum may determine the appearance of natural images.* Vision Res, 1993. **33**(1): p. 141-5.
14. van Hateren, J.H. and D.L. Ruderman, *Independent component analysis of natural image sequences yields spatio- temporal filters similar to simple cells in primary visual cortex.* Proc R Soc Lond B Biol Sci, 1998. **265**(1412): p. 2315-20.
15. van Hateren, J.H. and A. van der Schaaf, *Independent component filters of natural images compared with simple cells in primary visual cortex.* Proc Biol Sci, 1998. **265**(1394): p. 359-66.
16. Simoncelli, E.P., *Statistical modeling of photographic images*, in *Handbook of Image and Video Processing*, A.C. Bovic, Editor 2005, Academic Press: Burlington, MA. p. 431-441.
17. Lyu, S. and E.P. Simoncelli, *Nonlinear extraction of independent components of natural images using radial gaussianization.* Neural Comput, 2009. **21**(6): p. 1485-519.
18. Zetzsche, C. and U. Nuding, *Nonlinear and higher-order approaches to the encoding of natural scenes.* Network, 2005. **16**(2-3): p. 191-221.
19. Martens, J.-B., *The Hermite Transform -- Applications.* IEEE Transactions on Acoustics Speech and Signal Processing, 1990. **38**(9): p. 1607-1618.
20. Martens, J.-B., *The Hermite Transform --Theory.* Ieee Transactions on Acoustics Speech and Signal Processing, 1990. **38**(9): p. 1595-1606.
21. Martens, J.B., *Local orientation analysis in images by means of the Hermite transform.* IEEE Transactions on Image Processing, 1997. **6**(8): p. 1103-1116.
22. vanDijk, A.M. and J.B. Martens, *Representation and compression with steered Hermite transforms.* Signal Processing, 1997. **56**(1): p. 1-16.
23. Refregier, A., *Shapelets - I. A method for image analysis.* Monthly Notices of the Royal Astronomical Society, 2003. **338**(1): p. 35-47.
24. Silvan-Cardenas, J.L. and B. Escalante-Ramirez, *The multiscale hermite transform for local orientation analysis.* Ieee Transactions on Image Processing, 2006. **15**(5): p. 1236-1253.
25. Victor, J.D. and B.W. Knight, *Simultaneously band and space limited functions in two dimensions, and receptive fields of visual neurons*, in *Springer Applied Mathematical Sciences Series*, E.Kaplan, J. Marsden, and K.R. Sreenivasan, Editors. 2003, Springer. p. 375-420.
26. Slepian, D. and H. Pollack, *Prolate spheroidal wave functions, Fourier analysis and uncertainty -I.* Bell Syst Tech J, 1961. **40**: p. 43-64.
27. Slepian, D., *Prolate spheroidal wave functions, Fourier analysis and uncertainty-IV:Extensions to many dimensions; generalized prolate spheroidal functions.* Bell Syst Tech, 1964. **43**: p. 3009-3057.
28. Knight, B. and L. Sirovich, *The Wigner transform and some exact properties of linear operators.* SIAM J Appl Math, 1982. **42**: p. 378-389.
29. Ruderman, D.L., *The statistics of natural images.* Network Comput Neural Sys, 1994. **5**: p. 517-548.
30. Victor, J.D., et al., *Responses of V1 neurons to two-dimensional Hermite functions.* J Neurophysiol, 2006. **95**(1): p. 379-400.

31. Sharpee, T.O. and J.D. Victor, *Contextual modulation of V1 receptive fields depends on their spatial symmetry.* J Comput Neurosci, 2009. **26**(2): p. 203-18.
32. Sinz, F.H., E. Simoncelli, and M. Bethge, *Hierarchical modeling of local image features through L_p-nested symmetric distributions.* Advances in Neural Information Processing Systems, 2010. **22**: p. accepted.
33. Bethge, M., *Factorial coding of natural images: how effective are linear models in removing higher-order dependencies?* J Opt Soc Am A Opt Image Sci Vis, 2006. **23**(6): p. 1253-68.
34. Zhang, X. and S. Lyu, *Using projection kurtosis concentration of natural images for blind noise covariance matrix estimation.* IEEE Conference on Computer Vision and Pattern Recognition, 2014: p. DOI: 10.1109/CVPR.2014.367.
35. Motoyoshi, I., et al., *Image statistics and the perception of surface qualities.* Nature, 2007. **447**(7141): p. 206-9.
36. Graham, D., et al., *Preference for luminance histogram regularities in natural scenes.* Vision Res, 2016. **120**: p. 11-21.
37. Portilla, J. and E.P. Simoncelli, *A parametric texture model based on joint statistics of complex wavelet coefficients.* International Journal of Computer Vision, 2000. **40**(1): p. 49-71.

Figures for: Hu and Victor,

Two-Dimensional Hermite Filters Simplify the Description of High-Order Statistics of Natural Images

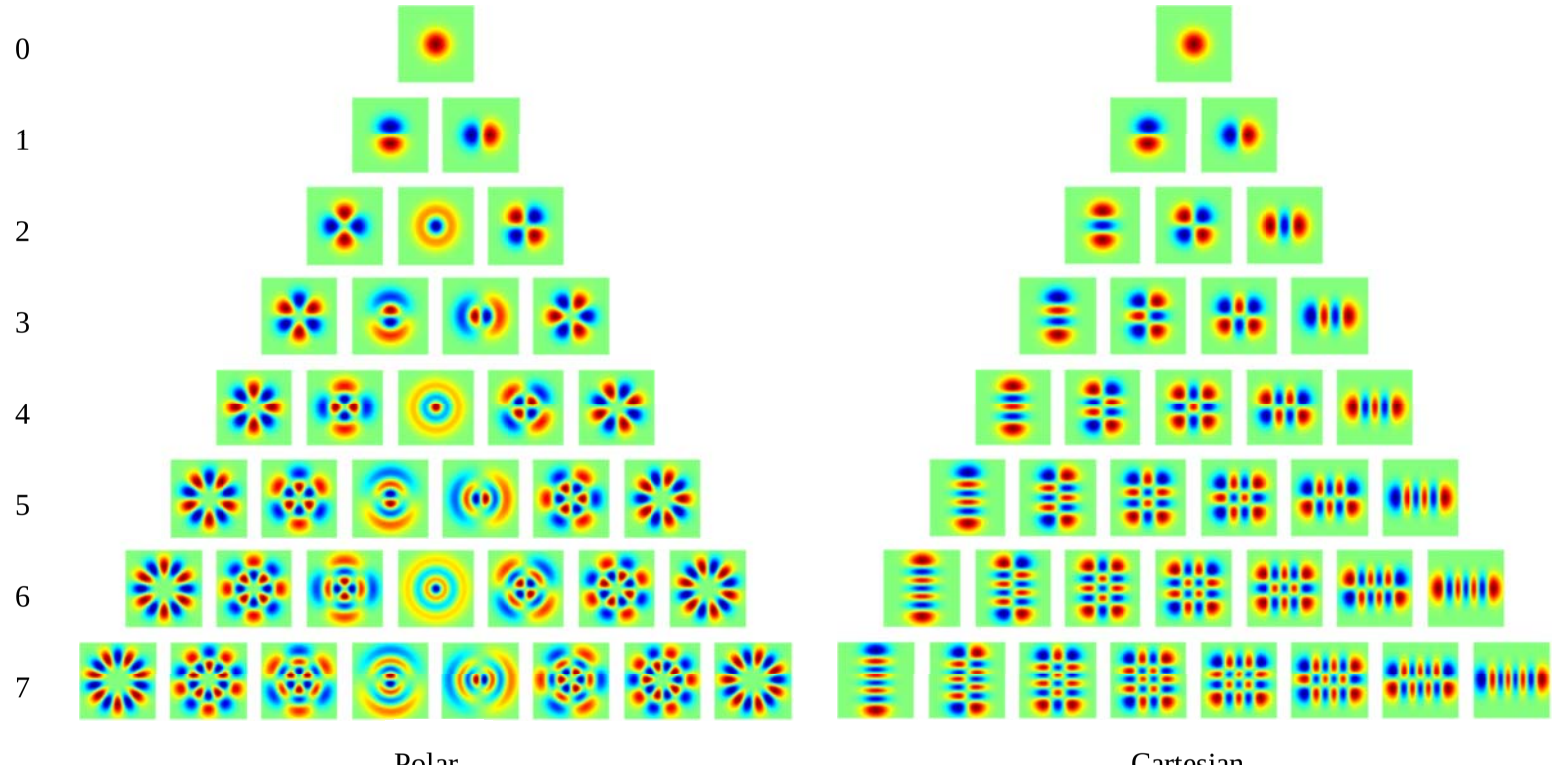

Figure 1. Two-dimensional Hermite (TDH) functions of rank 0 to 7, in (A) polar form and (B) Cartesian form. The pseudocolor scale (red positive, blue negative) is chosen separately for each function to cover the entire range. (Modified from Figure 1 in [30], Victor et al. 2006, J. Neurophysiol., Am Physiol Soc, with permission.)

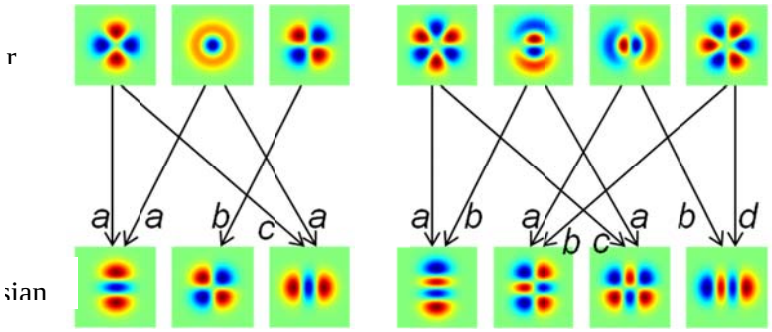

Figure 2. Cartesian TDH functions are linear combination of polar TDH functions. Examples are shown for rank 2 (left) and rank 3 (right). For rank 2, the coefficients are $a = \sqrt{2}/2, b = 1, c = -\sqrt{2}/2$. For rank 3, the coefficients are $a = 1/2, b = \sqrt{3}/2,$ $c = -\sqrt{3}/2, d = -1/2$.

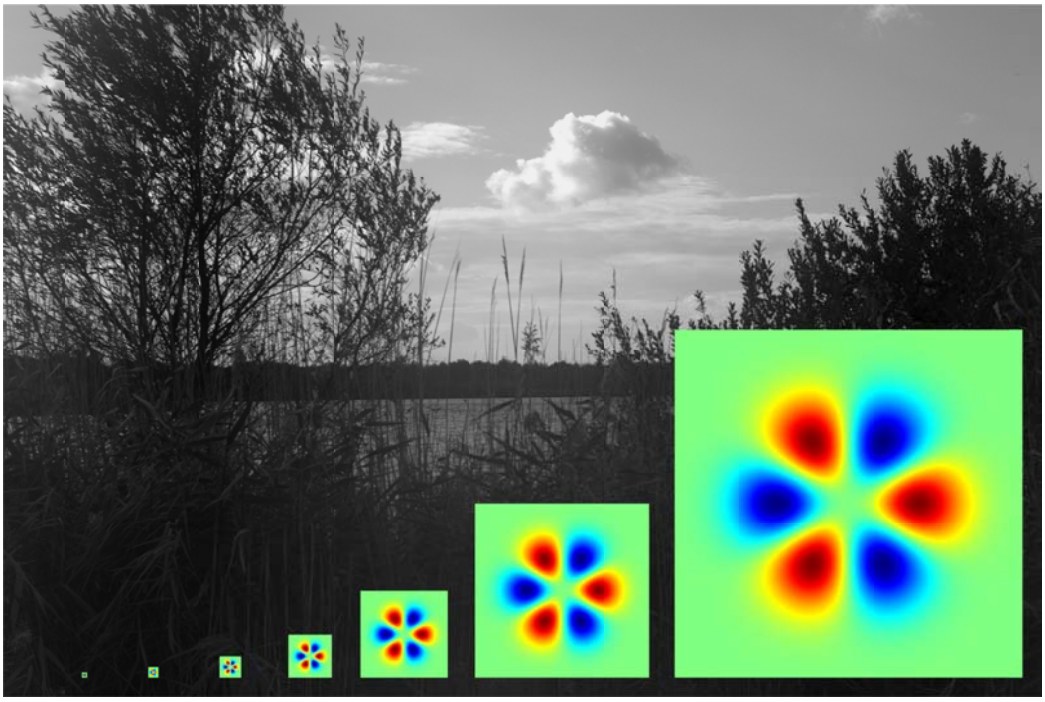

Figure 3. The seven filter sizes used to calculate image statistics, compared to the size of natural images used in this study (1536×1024).

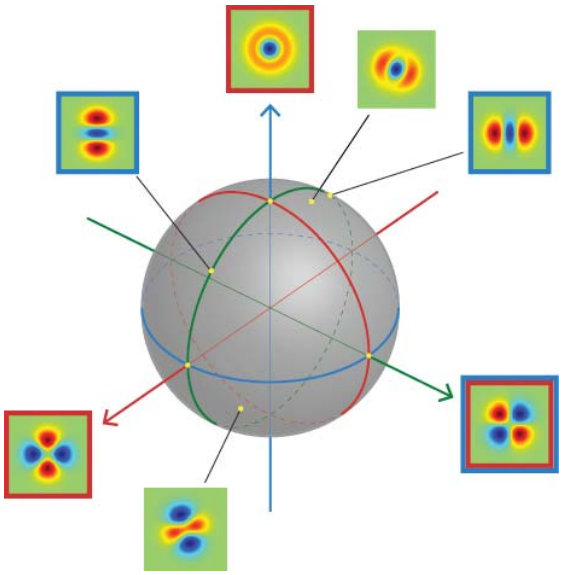

Figure 4. Generalized steerability of the rank 2 TDH filters. Each unit-magnitude filter corresponds to a point on the surface of a sphere. The polar and Cartesian basis functions form two sets of orthogonal coordinate axes. Filters with a red frame are polar TDH filters; filters with blue frame are Cartesian TDH filters; one filter is in both sets as indicated by its two frames. Filters without frame are intermediate filters; they can be constructed from a linear combination of either polar or Cartesian filters.

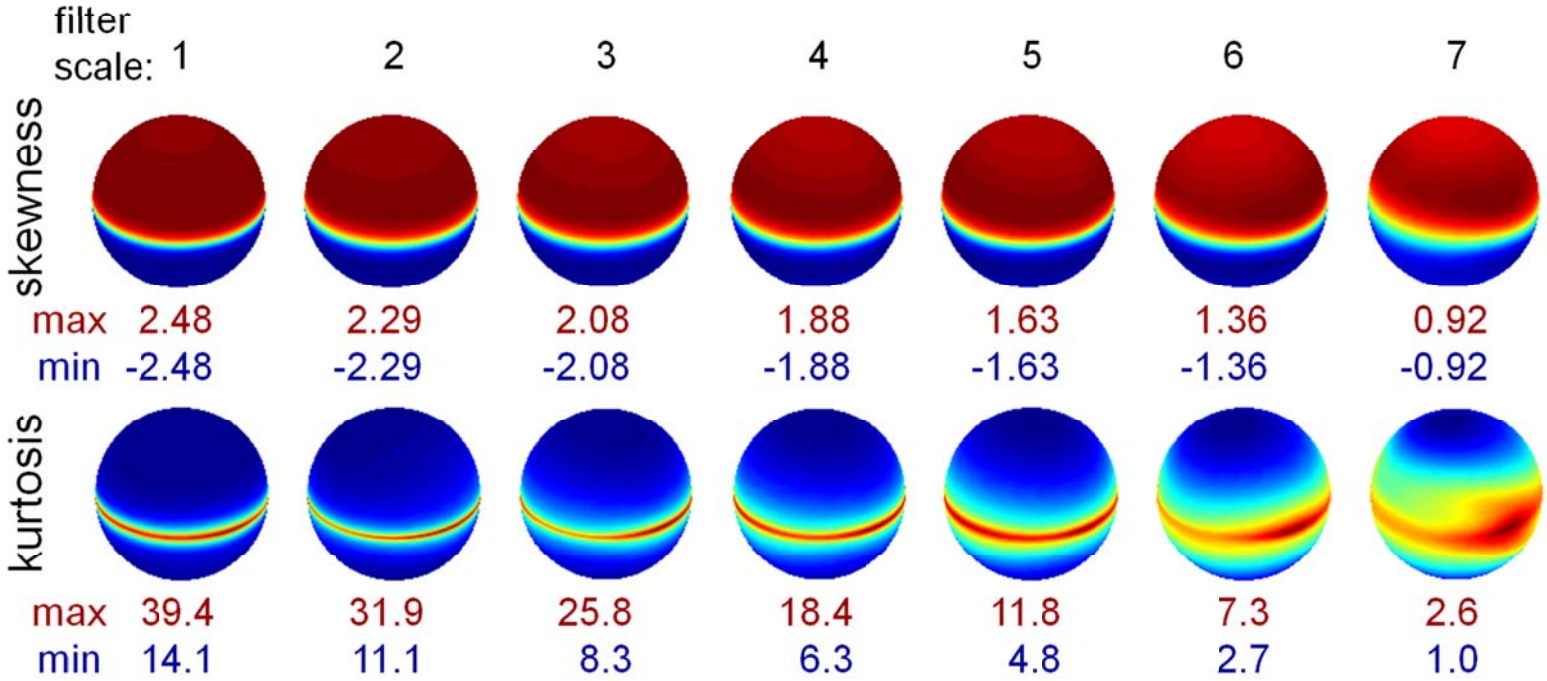

Figure 5. Skewness and kurtosis for natural images filtered by rank 2 TDH filters across 7 spatial scales. Each sphere represents the filter space of unit-length rank 2 TDH filters (oriented as shown in Figure4). Skewness and kurtosis are averaged across all filtered images, and plotted as a function of direction in the filter space. The pseudocolor scales for each skewness and kurtosis map are set to range from blue (minimum) to red (maximum). The minimum and maximum skewness and kurtosis values are shown under each sphere.

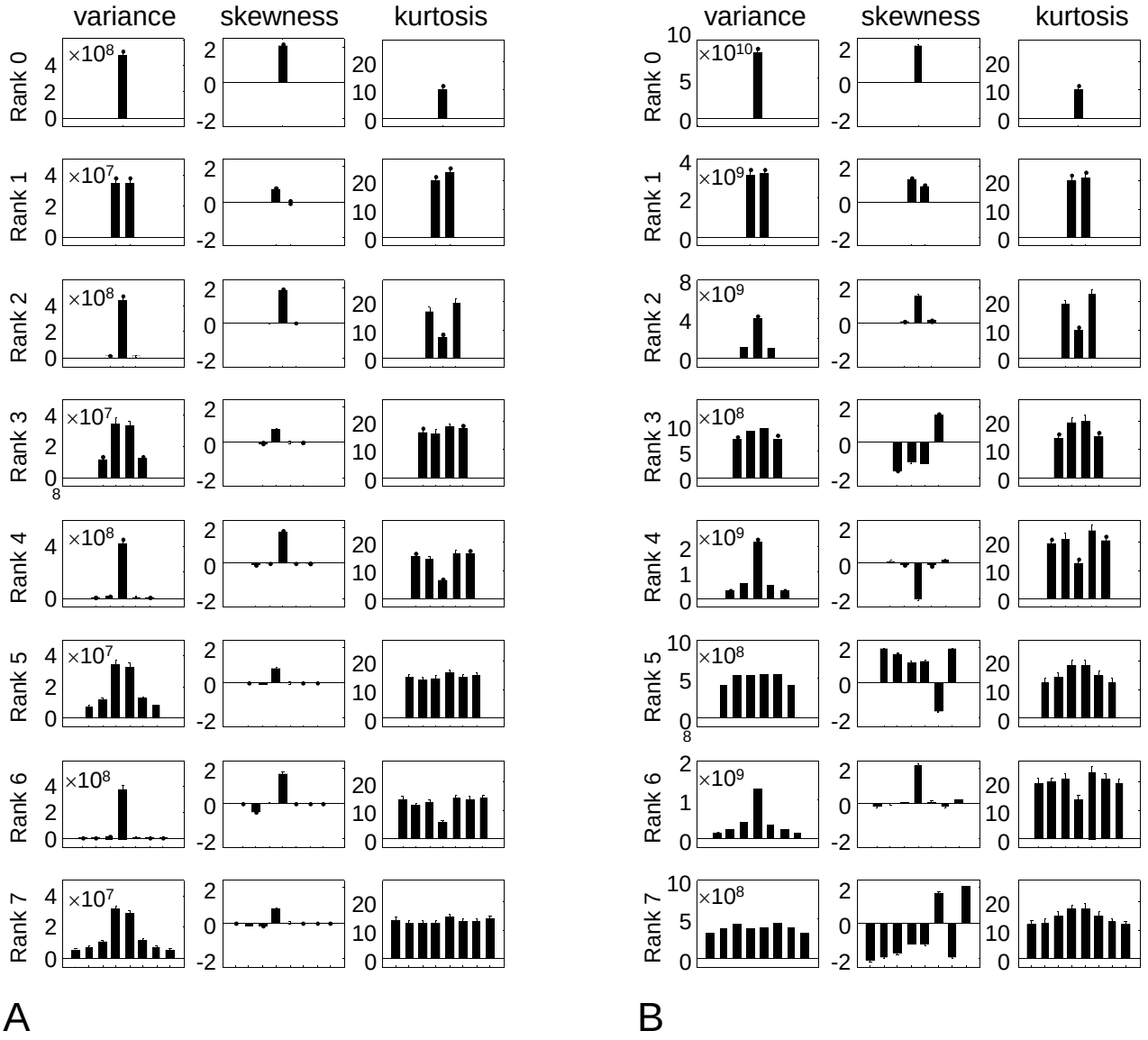

Figure 6. Variance, skewness, and kurtosis for (A) natural images filtered by polar TDH filters of rank 0 to 7 (spatial scale 4), and (B) modified TDH filters in which the polynomial component is replaced by its sign. Error bars are 3 SEM.

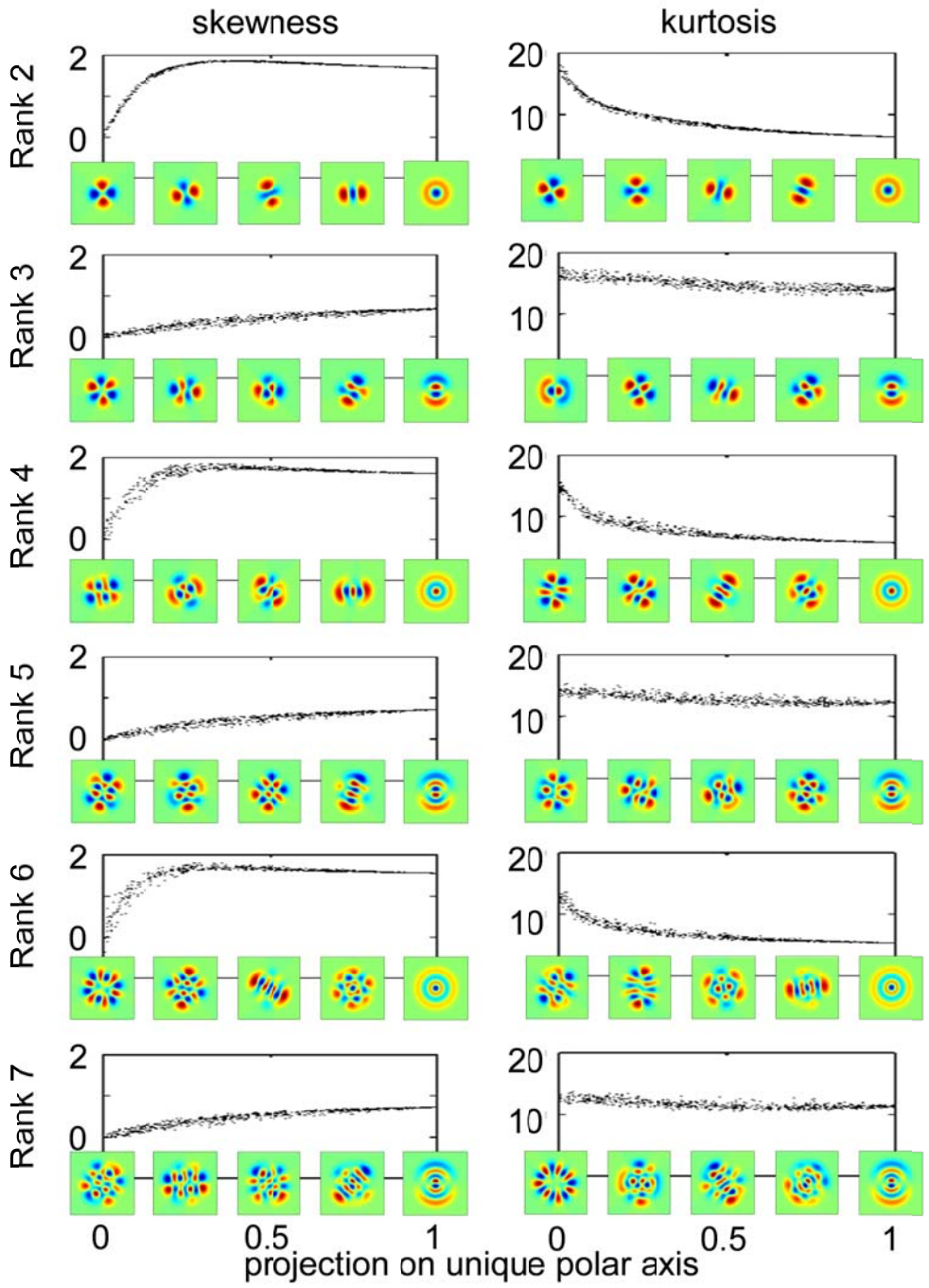

Figure 7. Skewness and kurtosis of natural images filtered by 1000 random TDH filters of rank 2 to 7, at scale 4. The abscissa is the projection of each random TDH filter onto the polar TDH filter shown at the lower right of each plot, which is the target-like filter for even ranks and the filter with a single, horizontal inversion axis for odd ranks. The filters placed along the abscissa are examples of filters whose projections onto the rightmost polar filter are 0, 0.25, 0.5, and 0.75. They illustrate the diversity of filters with a given value of the projection; the examples shown for the skewness and kurtosis columns at corresponding points along the abscissa are interchangeable.

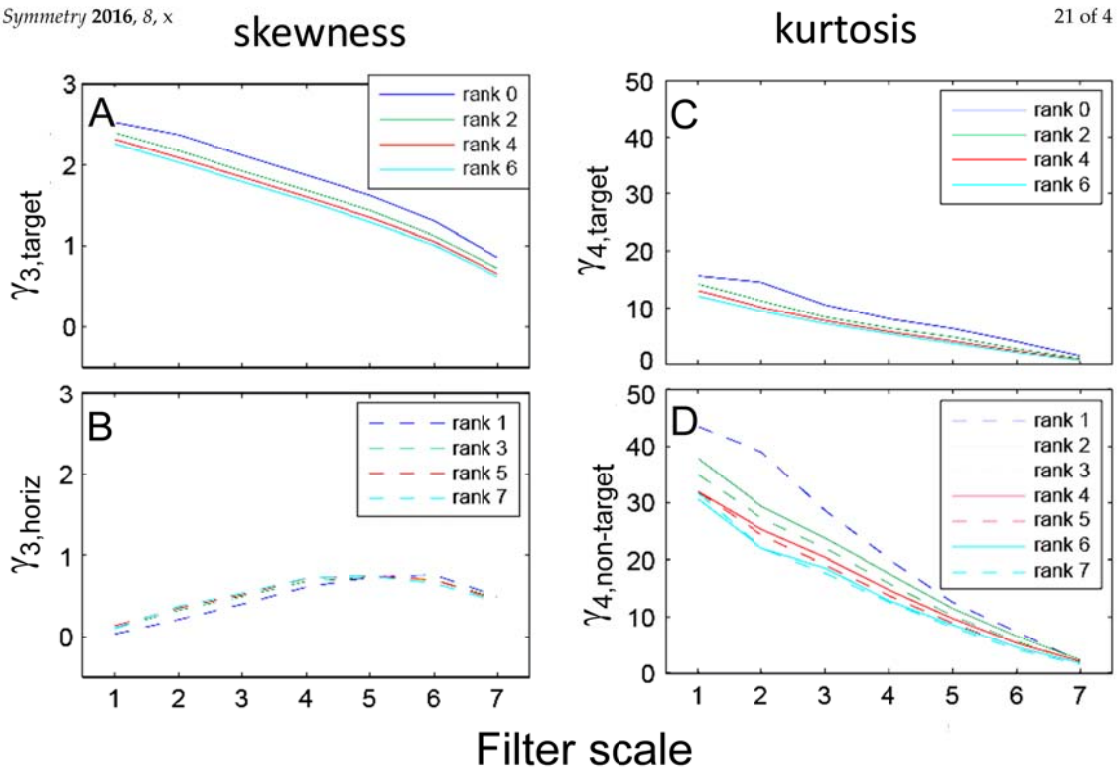

Figure 8. At each spatial scale, skewness of TDH-filtered images is characterized by two values: $\gamma_{3,target}$ (A) for even ranks and $\gamma_{3,horiz}$ for odd ranks (B), and kurtosis is characterized by $\gamma_{4,target}$ (C) for even ranks and $\gamma_{4,non-target}$ for all ranks (D).

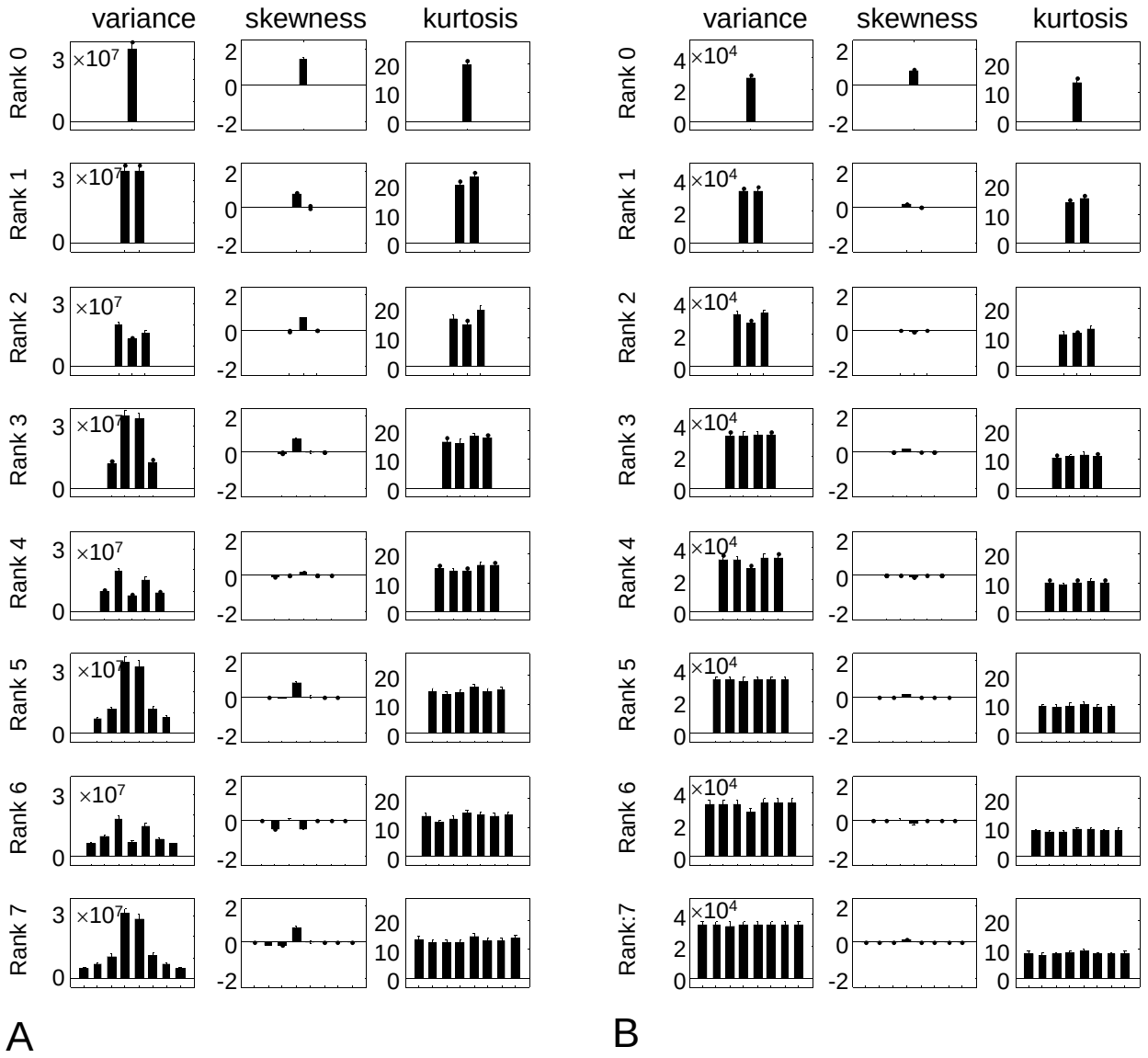

Figure 9. Variance, skewness, and kurtosis for (A) natural images filtered by polar TDH filters of rank 0 to 7 (spatial scale 4) after local mean subtraction. (B) as in (A), but natural images are whitened prior to analysis. Error bars are 3 SEM.

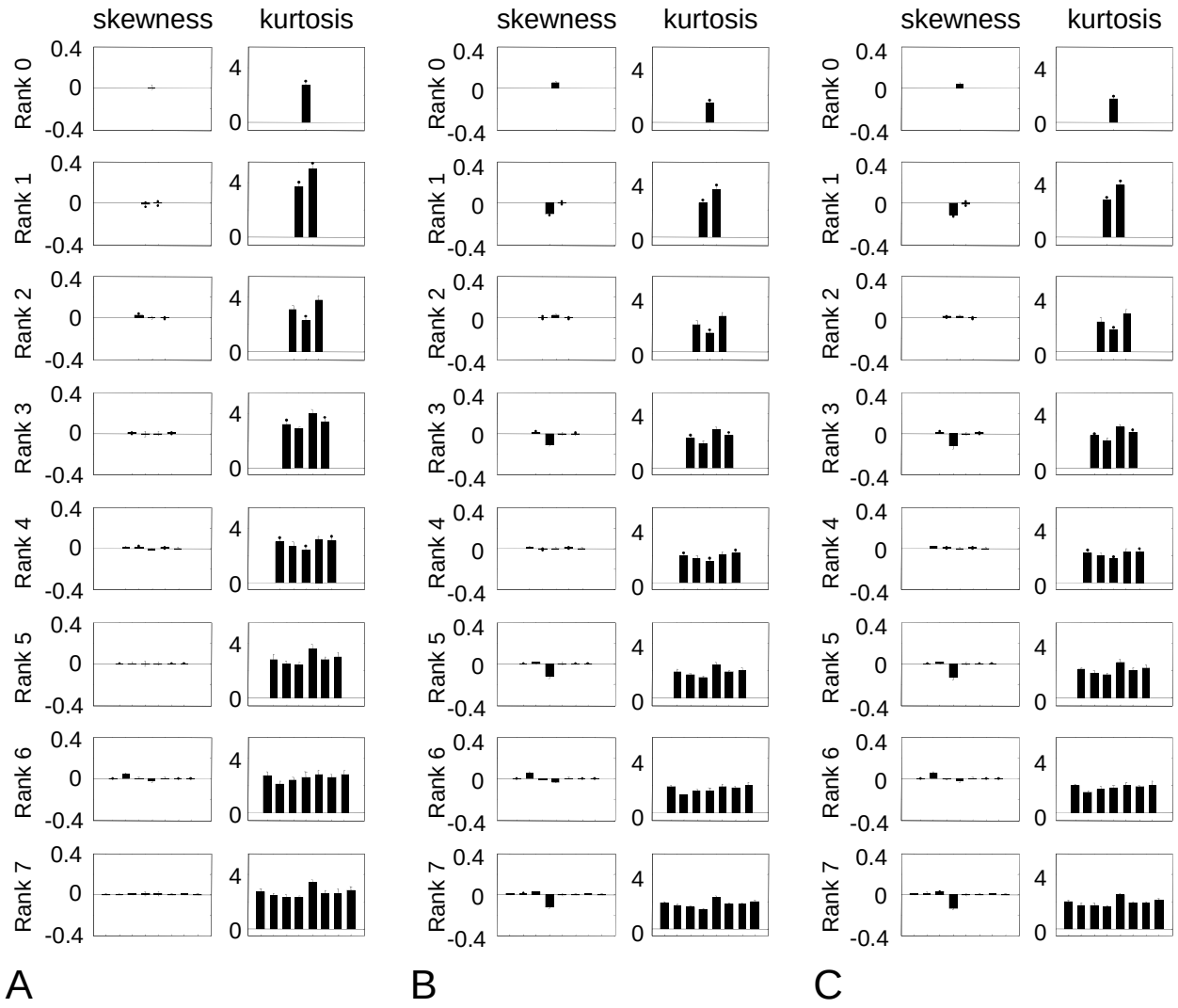

Figure 10. Skewness, and kurtosis TDH filters of rank 9 to 7 (spatial scale 4) processed by pointwise nonlinearities prior to analysis. (A) logarithmic transformation, (B) histogram equalization, (C) Gaussian luminance distribution. Error bars are 3 SEM.