

Duality of Log-Polar Image Representations in the Space and Spatial-Frequency Domains

Antonio Taberero, Javier Portilla, and Rafael Navarro

Abstract— In this paper, we study the result of applying a lowpass variant filtering using scaling-rotating kernels to both the spatial and spatial-frequency representations of a two-dimensional (2-D) signal (image). It is shown that if we apply this transformation to a Fourier pair, the two resulting signals can also form a Fourier pair when the filters used in each domain maintain a *dual* relationship. For a large class of “self-dual” filters, a perfect symmetry exists, so that the lowpass scaling-rotating variant filtering (SRVF) is the same in both domains, thus commuting with the Fourier transform operator. The lowpass SRVF of an image is often referred to as a “foveated” image, whereas its Fourier pair (the lowpass SRVF of its spectrum) can be realized as a local spectrum estimation around the point of attention. This lowpass SRVF is equivalent to a log-polar warping of the image representation followed by a lowpass invariant filtering and the corresponding inverse warping. The use of the log-polar warped representation allows us to extend the one-dimensional (1-D) scale transform to higher dimensions, in particular to images, for which we have defined a scale-rotation invariant representation. We also present an efficient implementation using steerable filters to compute both the foveated image and the local spectrum.

Index Terms— Foveated vision, local spectrum, log-polar mapping, scaling-rotating filters, scale transform, steerable filters, variant filtering, wavelets.

I. INTRODUCTION

LOG-POLAR representations play an important role in image processing and analysis. In the spatial domain, log-polar schemes [1] have been used to model the strongly inhomogeneous sampling of the retinal image by the human visual system. The number of receptor units (either at the retina [2] or in the later cortical stages [3]) drops rapidly with eccentricity while their size increases, causing a fast decay in visual acuity [4]. With this strategy, the visual system offers a tradeoff combining the advantages of a wide field of view with a high-resolution central area (fovea). Similar schemes have been introduced in artificial vision [5]–[7], where the use of “foveated” images helps reduce the amount of information to be processed.

Manuscript received October 13, 1997; revised February 26, 1999. This work was supported by the Spanish Comisión Interministerial de Ciencia y Tecnología (CICYT) under Grant TIC97–0325. The associate editor coordinating the review of this paper and approving it for publication was Prof. Moeness Amin.

A. Taberero is with the Facultad de Informática, Universidad Politécnica de Madrid, Madrid, Spain (e-mail: ant@asterix.fi.upm.es).

J. Portilla and R. Navarro are with the Instituto de Óptica, CSIC, Madrid, Spain.

Publisher Item Identifier S 1053-587X(99)06736-7.

In the spatial-frequency domain, methods such as short time Fourier analysis using an adaptive window [8], multiresolution local-spectrum estimation, or wavelet analysis [9], [10] usually involve a logarithmic sampling of the frequency domain. For images, a log-polar sampling is the natural extension, although separable log-log schemes are also used [11]. Again, the human visual system seems to apply a similar strategy for local spectrum estimation. Neurons with receptive fields having bandpass responses [12] are grouped in visual channels that have been modeled with banks of self-similar filters distributed in a log-polar grid [13]. The fast ($1/f^2$) decay of the power spectrum of natural images [14] suggests an additional reason for the use of filters whose bandwidth scales with frequency, resulting in an octave-distributed (log) scheme.

In practice, whether in the spatial or spatial-frequency domain, log-polar representations must eventually be sampled. The problem of preventing potential aliasing artifacts is then critical since the sampling interval becomes increasingly larger with eccentricity. Therefore, log-polar schemes demand a lowpass variant filtering in the linear domain, specifically using self-similar (scaled and rotated) variant kernels, to adapt locally the signal bandwidth to the varying sampling interval. From now on, we will be referring to this operation as a scaling-rotating variant filtering (SRVF).

As we have mentioned, log-polar representations have been extensively treated in the literature. However, little attention has been paid to the connection between log-polar representations in both the space and frequency domain, and to our knowledge, a direct mathematical analysis of this link has not been made. In this paper, we study such connection relating the signals resulting of applying a SRVF operator to an image and to its corresponding spatial-frequency spectrum.

In Section II, we show that the application of a lowpass variant filtering using scaled and rotated versions of a prototype filter in one domain is maintained as a similar SRVF (now scaling and rotating a *dual* filter) in the conjugated domain. We call this property *duality* of a SRVF operator in both the space and spatial-frequency domains. Although the actual filters applied need not be the same in both domains, a simple relation exists between a pair of dual filters. Moreover, there is a class of *self-dual* filters with the same expression in both domains, making the SRVF operator commute with the Fourier transform. While the result of lowpass SRVF in the spatial domain is interpreted as a log-polar “foveated” image, we show in Section III that a lowpass SRVF in the frequency domain provides an estimate of the local spectrum around

the point of fixation. In Section IV, we establish a straightforward connection with Cohen's one-dimensional (1-D) scale transform [15], proposing an extension of the scale transform to N dimensions, which could be useful to deal with scale and rotational invariances in images [16], [17]. Finally, in Section V, we address the problem of moving the attention point, presenting efficient implementations based on steerable filters to compute either the local spectrum or the foveated image at multiple fixation points.

II. DUALITY OF A SELF-SIMILAR LOWPASS VARIANT FILTERING

In order to prevent aliasing when sampling in a log-polar grid, we need to apply a lowpass variant filtering that is locally adapted to the variant sampling interval. In the log-polar case, this corresponds to using scaled and rotated (self-similar) filters. In this section, we show that a straightforward connection exists between such an SRVF when performed in the space (Section II-A) and the spatial-frequency domain (Section II-B). The following analysis is restricted to images (2-D), although it can be immediately extended to higher dimensions. A detailed analysis of the 1-D case can be found in [18].

A. Lowpass SRVF in the Spatial Domain

We depart from an image expressed either in polar coordinates $f(r, \theta)$ or, in vector notation, $f(\vec{r})$. In order to perform an SRVF, we first build a variant kernel H as a scaled and rotated version of a "prototype" function $P(r, \theta)$, maintaining a constant volume

$$H(\vec{r}, \vec{s}) = \frac{1}{r^2} P\left(\frac{s}{r}, \varphi - \theta\right). \quad (1)$$

Then, the SRVF can be expressed as the inner product between the image $f(\vec{r})$ and $H(\vec{r}, \vec{s})$

$$\begin{aligned} c(\vec{r}) &= \mathcal{L}_P(f(\vec{r})) = \langle f(\vec{r}), H^*(\vec{r}, \vec{s}) \rangle \\ &= \iint_{\mathcal{R}^2} d\vec{s} f(\vec{s}) H(\vec{r}, \vec{s}) \end{aligned} \quad (2)$$

where \mathcal{L}_P denotes the SRVF operator (\mathcal{L}) using P as its prototype filter. This filter $P(r, \theta)$ should have lowpass characteristics (its frequency response being centered around $\vec{\omega} = 0$) being located around $r = 1$, $\theta = 0$. In addition, the module of $P(r, \theta)$ should be integrable, that is

$$P(r, \theta) \in L^1(\mathcal{R}^2) \Leftrightarrow \|P\|_1 = \iint_{\mathcal{R}^2} d\vec{r} |P(\vec{r})| < \infty. \quad (3)$$

Although it is not necessary, for convenience, we will restrict ourselves to real prototype filters. Additional integral conditions on $P(r, \theta)$ will be discussed below.

Given the lowpass nature of the prototype filter $P(r, \theta)$, the filtered image

$$c(r, \theta) = \int_0^{2\pi} d\varphi \int_0^{2\pi} s ds f(s, \varphi) \frac{1}{r^2} P\left(\frac{s}{r}, \varphi - \theta\right) \quad (4)$$

consists of a weighted local average of the values of the original image around (r, θ) . Due to the scaling, the spatial

support of the variant kernel becomes larger when moving away from the origin, thus adjusting its lowpass effect to the increasing sampling interval of a log-polar grid.

An alternative approach to prevent aliasing in a log-polar grid consists of first warping the image so that the sampling grid becomes uniform and then applying an invariant lowpass filter in the warped (l_r, θ) domain, where $l_r = \log r$. Both approaches are equivalent if we define the following warping (preserving the energy) to take the original image into the log-warped domain:

$$\begin{aligned} f(r, \theta) & \quad r \in \mathcal{R}^+, \theta \in [0, 2\pi) \\ \hat{f}(l_r, \theta) &= e^{l_r} f(e^{l_r}, \theta) \quad l_r = \log r \in \mathcal{R}, \theta \in [0, 2\pi). \end{aligned} \quad (5)$$

By applying the above transformation to $c(r, \theta)$, $f(r, \theta)$, and $P(r, \theta)$ in (4), it follows that

$$\begin{aligned} \hat{c}(l_r, \theta) &= \int_0^{2\pi} d\varphi \int_{-\infty}^{\infty} dl_s \hat{f}(l_s, \varphi) \hat{P}(l_s - l_r, \varphi - \theta) \\ &= \hat{f}(l_r, \theta) * \hat{P}(-l_r, -\theta). \end{aligned} \quad (6)$$

Thus, the former SRVF is translated into a simple convolution with the filter $\hat{P}(-l_r, -\theta)$ in the warped domain (note that the integral in φ corresponds to a circular convolution). The filter used (\hat{P}) corresponds to the warped version of the former prototype function P . It is also a lowpass filter that is now centered around the origin ($l_r = 0, \theta = 0$) as $P(r, \theta)$ peaked at $(1, 0)$.

B. Duality of a SRVF in the Spatial and Spatial-Frequency Domains

In the previous section, we defined a lowpass SRVF operator (\mathcal{L}_P) acting on an image. Let us now study the Fourier transform of the filtered image $c(\vec{r}) = \mathcal{L}_P\{f(\vec{r})\}$. If we Fourier transform $c(\vec{r})$ in (2) and express the original image $f(\vec{s})$ as a function of its spectrum $F(\vec{\omega})$, we obtain

$$\begin{aligned} C(\vec{\omega}) &= \frac{1}{2\pi} \iint_{\mathcal{R}^2} d\vec{r} e^{-i\vec{\omega}\vec{r}} \left\{ \iint_{\mathcal{R}^2} d\vec{s} H(\vec{r}, \vec{s}) \right. \\ &\quad \left. \cdot \left(\frac{1}{2\pi} \iint_{\mathcal{R}^2} d\vec{\lambda} F(\vec{\lambda}) e^{i\vec{\lambda}\vec{s}} \right) \right\}. \end{aligned} \quad (7)$$

By regrouping terms in the above equation, we find that the relationship between the spectra $F(\vec{\lambda})$ and $C(\vec{\omega})$ is similar to that of the images $f(\vec{r})$ and $c(\vec{r})$ in (2)

$$\begin{aligned} C(\vec{\omega}) &= \langle F(\vec{\lambda}), \tilde{H}^*(\vec{\omega}, \vec{\lambda}) \rangle \\ &= \iint_{\mathcal{R}^2} d\vec{\lambda} F(\vec{\lambda}) \tilde{H}(\vec{\omega}, \vec{\lambda}). \end{aligned} \quad (8)$$

Thus, a spatial-variant filtering using $H(\vec{r}, \vec{s})$ is maintained as another variant filtering in the spatial-frequency domain, now with a different kernel $\tilde{H}(\vec{\omega}, \vec{\lambda})$. By comparing (7) and (8), we see that both kernels are related through a double 2-D Fourier transform, plus an axis reflection

$$\tilde{H}(\vec{\omega}, \vec{\lambda}) = \left(\frac{1}{2\pi} \right)^2 \iint_{\mathcal{R}^2} d\vec{r} e^{-i\vec{\omega}\vec{r}} \iint_{\mathcal{R}^2} d\vec{s} e^{i\vec{\lambda}\vec{s}} H(\vec{r}, \vec{s}). \quad (9)$$

Although the above relationship holds for any variant filtering, the case of a self-similar SRVF presents a particularly interesting property. When the variant spatial kernel $H(\bar{r}, \bar{s})$ comes from scaling and rotating a prototype filter (1) $P(r, \theta)$, the expression of the new variant kernel in the frequency domain can be obtained through (9) as

$$\tilde{H}(\bar{\omega}, \bar{\lambda}) = \frac{1}{\lambda^2} P\left(\frac{\omega}{\lambda}, \Phi - \Omega\right) \quad (10)$$

where (ω, Ω) and (λ, Φ) are the polar coordinates of $\bar{\omega}$ and $\bar{\lambda}$, respectively. Therefore, the variant kernel operating in the frequency domain has the same expression as its spatial counterpart, except for the role of the dummy radial variables (λ in frequency and s in space) that has been interchanged. Furthermore, if we define a *dual* filter \tilde{P} as

$$\tilde{P}(r, \theta) = \frac{1}{r^2} P\left(\frac{1}{r}, \theta\right) \quad (11)$$

the new spatial-frequency variant kernel is also a scaled and rotated version of the dual filter

$$\tilde{H}(\bar{\omega}, \bar{\lambda}) = \frac{1}{\omega^2} \tilde{P}\left(\frac{\lambda}{\omega}, \Phi - \Omega\right). \quad (12)$$

Consequently, an SRVF operator (\mathcal{L}) using a prototype filter P in one domain (\mathcal{L}_P) is maintained as a similar SRVF (\mathcal{L}) in the conjugate domain but now using the dual filter \tilde{P} ($\mathcal{L}_{\tilde{P}}$):

$$\begin{aligned} C(\bar{\omega}) &= \iint_{\mathcal{R}^2} d\bar{\lambda} F(\bar{\lambda}) \frac{1}{\lambda^2} P\left(\frac{\omega}{\lambda}, \Phi - \Omega\right) \\ &= \iint_{\mathcal{R}^2} d\bar{\lambda} F(\bar{\lambda}) \frac{1}{\omega^2} \tilde{P}\left(\frac{\lambda}{\omega}, \Phi - \Omega\right) \\ &= \mathcal{L}_{\tilde{P}}\{F(\bar{\omega})\}. \end{aligned} \quad (13)$$

We call this property *duality* of a SRVF in both the spatial and spectral domains, which is expressed as

$$\mathcal{F}\{\mathcal{L}_P\{f(r, \theta)\}\} = \mathcal{L}_{\tilde{P}}\{\mathcal{F}\{f(r, \theta)\}\} \quad (14)$$

where \mathcal{F} denotes a Fourier transform and \mathcal{L} the SRVF operator using either P (\mathcal{L}_P) or \tilde{P} ($\mathcal{L}_{\tilde{P}}$).

Finally, we will show how the resulting filtered spectrum $C(\bar{\omega})$ in (13) can also be obtained in the log-polar warped domain. By applying the same log-polar warping (5) to the spatial-frequency variables to each function in (13), we obtain

$$\hat{C}(l_w, \Omega) = \int_0^{2\pi} d\Phi \int_{-\infty}^{\infty} d(l_\lambda) \hat{F}(l_\lambda, \phi) \hat{P}(l_w - l_\lambda, \phi - \Omega). \quad (15)$$

The only difference between the log-polar frequency (15) and space (6) domains is that the convolving filter is now $\hat{P}(l_w, -\Omega)$ instead of $\hat{P}(-l_r, -\theta)$. Therefore, in the log-polar warped domain, the relationship between a pair of dual filters is a simple one-axis mirror reflection.

C. Additional Integrability Conditions on $P(r, \theta)$

In the previous paragraphs, we have implicitly assumed that the prototype filter and its dual counterpart are such that both (4) and (13) are well defined. However, it might be possible that for a given $P(r, \theta)$, the Fourier transform of the resulting $c(\bar{r})$ were not defined. Furthermore, even if $C(\bar{\omega})$ were defined, it could happen that its inverse Fourier transform did not exist. We avoid these potential problems by imposing some additional constraints on the prototype filter.

Let us suppose that both the original image $f(\bar{r})$ and its spectrum $F(\bar{\omega})$ belong to the space of finite energy signals $L^2(\mathcal{R}^2)$. Then, we have to find a condition on $P(r, \theta)$, ensuring that both $c(\bar{r})$ and $C(\bar{\omega})$ remain in $L^2(\mathcal{R}^2)$ to guarantee that they also form a well-defined Fourier pair.

As the logarithmic warping in (5) preserves the energy (L^2 metric), it suffices to prove that both warped signals $\hat{c}(l_r, \theta)$ and $\hat{C}(l_w, \Omega)$ have finite energy. These signals are the result of a convolution between a finite energy signal [either $\hat{f}(l_r, \theta)$ or $\hat{F}(l_w, \Omega)$] and the lowpass filter $\hat{P}(l_r, \theta)$. Since the convolution between $f \in L^2$ and $g \in L^1$ remains in L^2 [19], a sufficient condition for both $\hat{c}(l_r, \theta)$ and $\hat{C}(l_w, \Omega)$ being in L^2 is that $\hat{P}(l_r, \theta) \in L^1$, that is

$$\begin{aligned} &\int_0^{2\pi} \int_{-\infty}^{\infty} d\theta dl_r |\hat{P}(l_r, \theta)| \\ &= \int_0^{2\pi} \int_0^{\infty} d\theta dr |P(r, \theta)| \\ &= \int_0^{2\pi} \int_0^{\infty} r dr d\theta \frac{|P(r, \theta)|}{r} \\ &= \iint_{\mathcal{R}^2} d\bar{r} \frac{|P(\bar{r})|}{|\bar{r}|} < \infty. \end{aligned} \quad (16)$$

By dividing this integral into two parts corresponding to $r \leq 1$ and $r > 1$, we find that

$$\begin{aligned} \iint_{\mathcal{R}^2} d\bar{r} \frac{|P(\bar{r})|}{r} &< \iint_{r \leq 1} d\bar{r} \frac{|P(\bar{r})|}{r^2} + \iint_{r > 1} d\bar{r} |P(\bar{r})| \\ &< \iint_{\mathcal{R}^2} d\bar{r} \frac{|P(\bar{r})|}{r^2} + \iint_{\mathcal{R}^2} d\bar{r} |P(\bar{r})| \\ &= \|\tilde{P}\|_1 + \|P\|_1. \end{aligned} \quad (17)$$

Consequently, a sufficient condition for (16) to hold is to impose that the dual filter \tilde{P} also belongs to $L^1(\mathcal{R}^2)$, that is

$$\|\tilde{P}\|_1 = \iint_{\mathcal{R}^2} d\bar{r} |\tilde{P}(\bar{r})| = \iint_{\mathcal{R}^2} d\bar{r} \frac{|P(\bar{r})|}{|\bar{r}|^2} < \infty. \quad (18)$$

Conditions (3) and (18) indicate a sufficient fast decay of $P(r, \theta)$ at both the origin and infinity. Although less restrictive conditions than having the dual filter belong to L^1 could possibly be found to ensure that $c(\bar{r})$ and $C(\bar{\omega})$ form a Fourier pair, (18) is especially appropriate here as it reinforces the symmetry between the spatial and spatial-frequency domains.

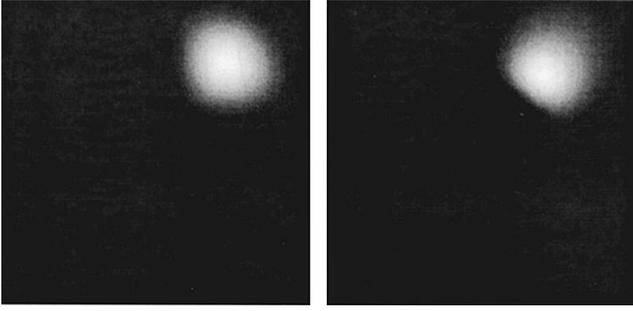


Fig. 1. Prototype filter $P(r, \theta)$ (left) as defined in (19) and its dual (20) counterpart \tilde{P} (right).

D. Examples and Interpretation

In order to illustrate the above concepts, let us consider a prototype function defined by

$$P(r, \theta) = r \exp(-A(r-1)^2) \exp(-B\theta^2) \quad (19)$$

corresponding to a real polar-separable lowpass integrating kernel with a maximum around $(r=1, \theta=0)$ and belonging to L^1 , as indicated. Parameters A and B determine its width and aspect ratio. This function (for $A=B=10$) is shown in the left panel of Fig. 1. Its corresponding dual filter, which is given by

$$\begin{aligned} \tilde{P}(r, \theta) &= \frac{1}{r^2} P\left(\frac{1}{r}, \theta\right) \\ &= \frac{1}{r^3} \exp\left(-A\left(\frac{1}{r}-1\right)^2\right) \exp(-B\theta^2) \end{aligned} \quad (20)$$

is shown in the right panel. Although different in general, the dual filter \tilde{P} preserves the energy of the original prototype filter P and maintains its lowpass nature. This dual filter also belongs to L^1 , indicating that P also verifies (18).

Fig. 2 illustrates the duality of a lowpass SRVF in both domains. The left panels in Fig. 2 display a Fourier pair: the image $f(\vec{r})$ on top and its frequency spectrum $F(\vec{\omega})$ at the bottom. The upper right panel shows a lowpass SRVF of the image $c(\vec{r}) = \mathcal{L}_P\{f(\vec{r})\}$ using the filter defined above (19). We can observe how the center of the original image (the point of attention) is preserved by the SRVF, whereas due to the scaling, the lowpass effect increases as we move away from the origin. The lower right panel shows the Fourier transform of the variant filtered image ($C(\vec{\omega}) = \mathcal{F}\{c(\vec{r})\}$). It is clear that when compared with the original spectrum, fine details have been preserved around the origin (DC) while being progressively lost toward the periphery, showing that $C(\vec{\omega})$ is also a lowpass SRVF of the spectrum, now using the dual filter, $C(\vec{\omega}) = \mathcal{L}_{\tilde{P}}\{F(\vec{\omega})\}$. This is the main result of this section. The Fourier transform of a lowpass SRVF is also a lowpass SRVF of the original spectrum (14). In other words, the upper and lower panels correspond to Fourier pairs that are maintained when applying a lowpass SRVF (that is, moving from left to right). In these examples, we have used narrower filters ($A=B=100$) than those depicted in Fig. 1, in order to reduce the lowpass effect and obtain a larger fovea.

The progressive degradation of visual quality toward the periphery shown in the upper right panel of Fig. 2 is similar

to what happens in the human visual system [2], [3], where much of the resources are devoted to the central area of the visual field (fovea). Therefore, we will be referring to images such as the one shown in the upper right panel of Fig. 2 as “foveated” images and use the term “foveation process” to describe the lowpass SRVF of an image.

A desirable property of an ideal foveated image would be the preservation of the value of the original image at the origin, reflecting the intuitive idea that there should be no distortion at the fixation point. If we relate the values at the origin of both $c(\vec{r})$ and $f(\vec{r})$, we find that

$$\begin{aligned} c(\vec{0}) &= \iint_{\mathcal{R}^2} d\vec{\omega} C(\vec{\omega}) \\ &= \iint_{\mathcal{R}^2} d\vec{\lambda} F(\vec{\lambda}) \iint_{\mathcal{R}^2} d\vec{\omega} \tilde{H}(\vec{\omega}, \vec{\lambda}) \\ &= \iint_{\mathcal{R}^2} d\vec{\lambda} F(\vec{\lambda}) \iint_{\mathcal{R}^2} d\vec{x} P(\vec{x}) \\ &= I_P f(\vec{0}) \end{aligned} \quad (21)$$

where I_P is the volume integral of the prototype filter. Since P belongs to L^1 (3), it is always possible to normalize P (as long as $\|P\|_1 \neq 0$, which is ensured since P is a lowpass filter) so that $I_P = 1$, and thus, the origin is maintained in the foveated image $c(\vec{0}, 0) = f(\vec{0}, 0)$.

Another convenient property would be that the mean intensity level across the image should be maintained after the variant filtering, which is equivalent to imposing that the origin of the spectrum (the DC component) be preserved in the foveated spectrum. Considering the duality between both domains, and similarly to (21), we find that

$$\begin{aligned} C(\vec{0}) &= \iint_{\mathcal{R}^2} d\vec{r} c(\vec{r}) \\ &= \iint_{\mathcal{R}^2} d\vec{s} f(\vec{s}) \iint_{\mathcal{R}^2} d\vec{r} H(\vec{r}, \vec{s}) \\ &= \iint_{\mathcal{R}^2} d\vec{s} f(\vec{s}) \iint_{\mathcal{R}^2} d\vec{x} \tilde{P}(\vec{x}) \\ &= I_{\tilde{P}} F(\vec{0}). \end{aligned} \quad (22)$$

Then, if the dual filter is normalized in volume ($I_{\tilde{P}} = 1$), the DC component of the original image $F(\vec{0})$ will be preserved in the local spectrum. The volume integrability of the dual filter is again ensured through (18), indicating that $\tilde{P} \in L^1$. However, since the normalization of $\tilde{P}(r, \theta)$ brings an additional condition on $P(r, \theta)$, we cannot ensure that $I_{\tilde{P}} = 1$ once I_P has been normalized. That means that in general, it is not possible for a given prototype filter to preserve both the origin and the DC through a SRVF (unless $I_P = I_{\tilde{P}}$). For instance, with the prototype P defined in (19), we find (for $A=B=10$) that $I_P = 0.165$, whereas the volume of its dual filter (20) is different, namely, $I_{\tilde{P}} = 0.157$. Therefore, it is not possible to normalize both P and \tilde{P} simultaneously, although the difference will usually be small, as in this case.

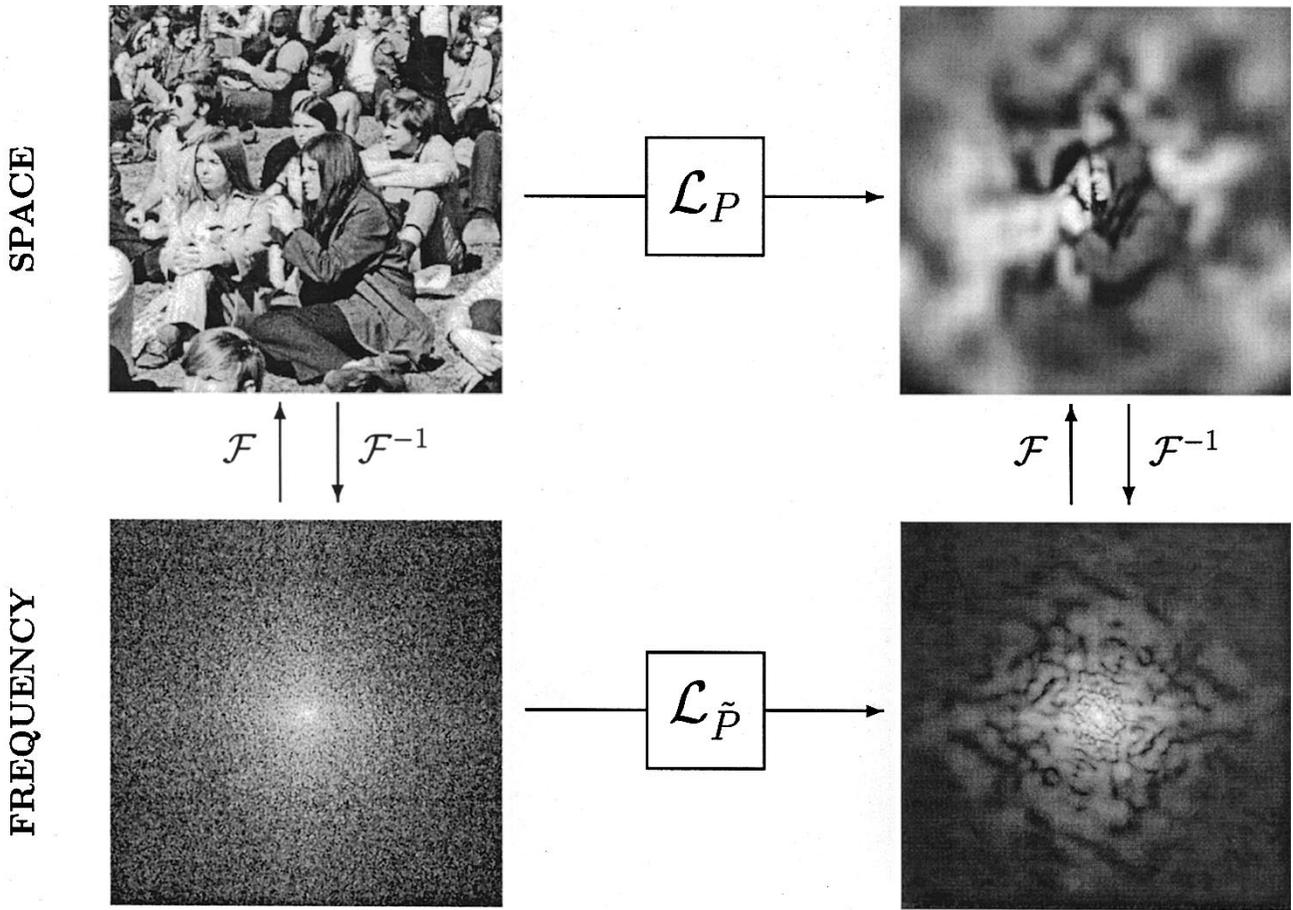


Fig. 2. Upper panel: Original image (left) and “foveated” image (right) obtained through a lowpass SRVF using the prototype filter P (\mathcal{L}_P). Lower panel: Moduli of the corresponding Fourier transforms (log-intensity scale) also related through a lowpass SRVF now using \tilde{P} ($\mathcal{L}_{\tilde{P}}$).

E. Self-Dual Filters

We have shown how an SRVF is similar in both the spatial (4) and frequency (13) domains as it relies on the same lowpass function $P(r, \theta)$. However, the differences between a filter and its dual counterpart \tilde{P} (Fig. 1) prevent a complete symmetry between domains. Nevertheless, it is straightforward to obtain *self-dual* kernels: a situation where both kernels of a dual pair are identical. The condition for self-duality is immediate in the log-polar domain. If (6) and (15) are to be equal, then the log-warped prototype filter must be even in its radial variable

$$\hat{P}(l_r, \theta) = \hat{P}(-l_r, \theta). \quad (23)$$

Consequently, given any 2-D function $Q(x, \theta)$ defined in $\mathcal{R}^+ \times [0, 2\pi)$, we can obtain a prototype filter generating self-dual kernels as

$$\begin{aligned} \hat{P}(l_r, \theta) &= Q(|l_r|, \theta) \\ P(r, \theta) &= \frac{1}{r} Q(|\log r|, \theta) = \tilde{P}(r, \theta) \end{aligned} \quad (24)$$

with the only condition (3) that $P(r, \theta) \in L^1$. Note that since $\tilde{P} = P$, the additional integrability condition (18) is trivially verified, and by normalizing $P(r, \theta)$ in volume, we accomplish $I_P = I_{\tilde{P}} = 1$ simultaneously. Therefore, another important motivation behind the use of self-dual filters is the preservation, after the lowpass SRVF, of both the signal at

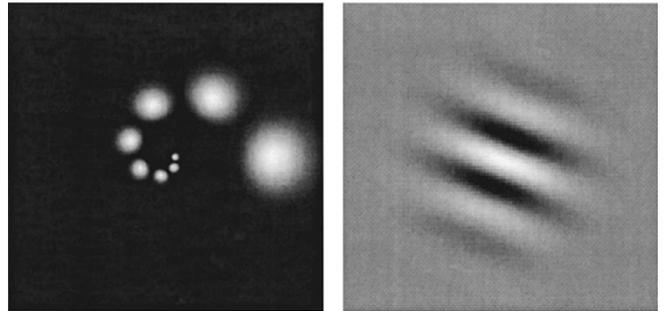


Fig. 3. Left: Several self-dual kernels for various scales and orientations generated using (25). Right: Inverse Fourier transform (real part) of one of them showing bandpass characteristics.

the point of attention and the DC component in the foveated image and spectrum, respectively.

Fig. 3 (left) shows several self-dual kernels (normalized in amplitude, instead of volume to improve visualization) obtained as scaled/rotated versions of a prototype filter generated from

$$Q(l_r, \theta) = \exp \{-(Al_r^2 + B\theta^2)\} \text{ with } A = B = 12. \quad (25)$$

The application of these kernels will also produce a progressive image degradation toward the periphery, but now, being self-dual, the “foveation” effect would be identical in the upper

and lower panels of Fig. 2. In fact, a lowpass SRVF using self-dual filters (\mathcal{L}_*) achieves perfect symmetry in both domains so that (14) becomes

$$\mathcal{F}\{\mathcal{L}_*\{f(r, \theta)\}\} = \mathcal{L}_*\{\mathcal{F}\{f(r, \theta)\}\} \quad (26)$$

making the Fourier transform operator commute with the lowpass SRVF operator.

As an interesting practical result, we present a simple method to obtain a self-dual filtering through a double application of any (non self-dual) filter P [18]. The key idea is to apply an SRVF using P to both the image (to obtain $c(\bar{r}) = \mathcal{L}_P\{f(\bar{r})\}$) and to its spectrum, obtaining another filtered image $c'(\bar{r}) = \mathcal{F}^{-1}\{\mathcal{L}_P\{F(\omega)\}\}$. Unless using self-dual filters, this new foveated image will differ from $c(\bar{r})$ since $c'(\bar{r})$ corresponds (14) to the application of the dual filter *directly* in the spatial domain:

$$\begin{aligned} c'(\bar{r}) &= \mathcal{F}^{-1}\{\mathcal{L}_P\{F(\omega)\}\} = \mathcal{F}^{-1}\{\mathcal{L}_P\{\mathcal{F}\{f(\bar{r})\}\}\} \\ &= \mathcal{F}^{-1}\{\mathcal{F}\{\mathcal{L}_{\tilde{P}}\{f(\bar{r})\}\}\} = \mathcal{L}_{\tilde{P}}\{f(\bar{r})\}. \end{aligned} \quad (27)$$

However, if we average $c(\bar{r})$ and $c'(\bar{r})$, the result is equivalent to applying the average of $\hat{P}(-l_r, -\theta)$ and $\hat{P}(l_r, -\theta)$ as the lowpass filter in the warped log-polar representation. This averaged filter is obviously symmetric in its first variable and, consequently, self-dual.

This procedure is of practical interest when the direct use of a self-dual filter is inconvenient. For instance, an implementation can be much more efficient using a prototype function with circular symmetry, which will not be self-dual. In that case, we can rely on the above procedure to obtain a self-dual filtered image or spectrum (at the cost of applying the filter twice).

III. FROM FOVEATED IMAGES TO LOCAL SPECTRUM ESTIMATION

In the previous section, we described a lowpass SRVF of an image as a stage of a foveation process (the other being the consequent log-polar sampling). Now, we turn our attention to the spatial-frequency domain. The Fourier transform of a foveated image or “foveated” spectrum (see Fig. 2) $C(\bar{\omega})$ admits a direct interpretation as a local spectrum estimation around the point of attention. This can be shown by applying Parseval’s identity to the right side of (13)

$$C(\omega, \Omega) = \int_0^{2\pi} \int_0^\infty r dr d\theta f(r, \theta) \tilde{p}(\omega r, \theta - \Omega) \quad (28)$$

where $\tilde{p}(r, \theta)$ is the inverse Fourier transform of the dual filter $\tilde{P}(\omega, \Omega)$, whose existence is ensured since $\tilde{P} \in L^1$. The right panel of Fig. 3 shows the real part of \tilde{p} , which is a bandpass-oriented filter spatially centered at the origin. Under these considerations, it becomes clear that $C(\omega, \Omega)$ in (28) is a local spectrum, i.e., the result of applying a (continuous) bank of analysis filters [rotated (Ω) and scaled (ω) versions of \tilde{p}] to an image to extract its frequency components (sub-bands) around a given point of fixation (the origin).

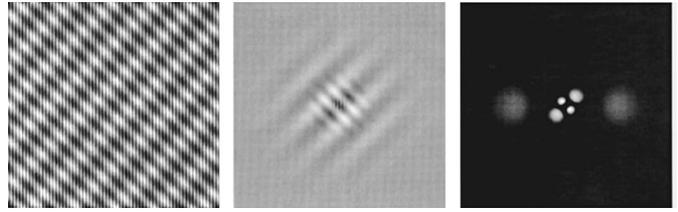


Fig. 4. Left: Original image containing three pure frequencies. Center: Foveated or SRVF image. Right: Fourier transform of the foveated image; the “foveated” or local spectrum.

Using scaled filters to analyze a signal is closely related to multiscale or wavelets methods [9], [10]. However, whereas a continuous wavelet transform collects scale information at every point simultaneously, with the subsequent redundancy, the local spectrum $C(\bar{\omega})$ gathers scale information only around the attention point. In order to having information about other positions, we must move the fixation point (see Section V). A more detailed discussion on the relationship between this “foveated” local spectrum and wavelets can be found in [18] for the 1-D case.

Fig. 4 further illustrates the interpretation of a foveated spectrum as a local spectrum estimate. The central panel shows the result of applying a lowpass SRVF to an image containing three pure frequencies (left). The highest frequency component disappears rapidly as we move away from the center, whereas the middle frequency is attenuated more gradually, and only the lowest frequency remains visible in the periphery. The Fourier transform of this foveated image is shown in the right panel. The three pure frequencies (delta functions in the original spectrum) have been progressively blurred, becoming “blobs” whose support increases with spatial frequency. Fig. 4 also shows how the Fourier transform of a foveated image $C(\bar{\omega}) = \mathcal{F}\{c(\bar{r})\}$ is closely related to other methods for local frequency estimation. For instance, sliding-window or short time Fourier analysis (STFA) methods [20] first select the local information by applying a spatial window and then compute its Fourier transform to estimate its frequency contents. The process here is essentially the same, except for the fact that we rely on a foveation, instead of a windowing, to extract the local information. As a result of foveation, the analysis window (Fig. 4 center) is automatically adapted to the present frequencies, as in other multiscale methods [9] or STFA methods with an adaptive window [8].

Alternatively, a lowpass SRVF of a global spectrum $C(\bar{\omega}) = \mathcal{L}\{F(\bar{\omega})\}$ also makes sense as a local spectrum estimate. The very low frequencies (which do not change much between a global or local spectral analysis due to their larger support) are relatively unaffected by the variant filtering. However, higher frequencies that are more likely to be modified when shifting from a global to a local perspective are progressively more affected by the lowpass SRVF.

IV. SCALE AND ROTATION INVARIANT REPRESENTATION OF IMAGES

In previous sections, we have shown that the lowpass SRVF needed to obtain either a foveated image or a local spectrum in the linear Cartesian space can be translated into a simple

convolution through a log-polar warping. Going one step further, that convolution becomes a simple product via the convolution theorem. In the 1-D case, applying the Fourier transform to the log-warped domain takes us to Cohen's scale transform [15]. We show next that it is immediate to generalize Cohen's transform to 2-D images and higher dimension signals.

Given an image f , we define $D_f(a, k)$ as the Fourier transform of its log-polar warping $\hat{f}(l_r, \theta)$:

$$D_f(a, k) = \frac{1}{2\pi} \int_0^{2\pi} d\theta e^{-ik\theta} \int_{-\infty}^{\infty} dl_r e^{-ial_r} \hat{f}(l_r, \theta) \quad (29)$$

with $k \in \mathcal{Z}, a \in \mathcal{R}$.

Note that the first variable a is continuous, whereas the second k is discrete due to the periodicity in θ . If we now express $D_f(a, k)$ as a function of the original image $f(r, \theta)$, we obtain

$$D_f(a, k) = \frac{1}{2\pi} \int_0^{2\pi} d\theta e^{-ik\theta} \int_0^{\infty} dr e^{-ial_r} f(r, \theta) \quad (30)$$

where $D_f(a, k)$ constitutes a generalization of Cohen's 1-D scale-transform [15] to the bidimensional case. In the same way that the modulus of the scale transform is invariant under scaling, the modulus of the above 2-D transform is invariant under both rotation and scaling of the image. This *scale-orientation* transform can be realized as the projection of the image $f(r, \theta)$ onto a set of basis functions $\gamma_{a,k}$:

$$D_f(a, k) = \langle f(\bar{r}), \gamma_{a,k}(\bar{r}) \rangle = \iint_{\mathcal{R}^2} d\bar{r} f(\bar{r}) \gamma_{a,k}^*(\bar{r})$$

with $\gamma_{a,k}(r, \theta) = \frac{1}{2\pi r} e^{i(al_r + k\theta)}$ (31)

whose radial part has already been proposed in [21]. Since the basis functions are orthonormal with respect to the defined inner product, the corresponding inverse transform is simply

$$f(r, \theta) = \langle D_f(a, k), \gamma_{a,k}^*(r, \theta) \rangle = \frac{1}{2\pi r} \sum_k e^{ik\theta} \int_{-\infty}^{\infty} da D_f(a, k) e^{ia \log r}. \quad (32)$$

Contrary to another alternative proposal to extend the scale transform to images [11], our transformation (30) separates the radial and angular dependencies in the signal by applying a 1-D scale transform only to the radial variable: the only one affected by a global change of scale. The Fourier analysis of the angular dependency is an additional feature through which rotational invariance is accomplished. The separation between the radial and angular variables allows us to extend the scale-orientation transform to N dimensions by incorporating into (31) the generalization of the radial basis functions proposed in [21].

Then, through the convolution theorem, the convolution in the warped log-polar spatial domain (6) becomes a multiplication in the new scale domain

$$D_c(a, k) = D_f(a, k) D_P(-a, -k) \quad (33)$$

where $D_c(a, k)$ and $D_P(a, k)$ are the *scale-orientation* trans-

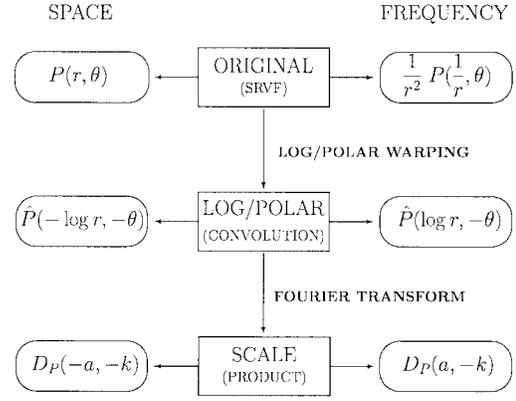


Fig. 5. Operations required to perform an SRVF when using a linear (upper), log-polar (middle), or scale (lower) representation in either the space (left) or spatial-frequency (right) domains.

forms of the foveated image and the prototype filter, respectively. Therefore, the foveation process is equivalent to filtering out the higher "scales" of the original image. The negative signs in (33) are simply a consequence of our using $\hat{P}(-l_r, -\theta)$ as the lowpass filter in the warped domain.

Fig. 5 summarizes the results presented so far in the different domains from the variant filtering using rotated/scaled filters that takes place in the linear domain (upper level) through the convolution in the log-polar warped domain (middle level) to end up with a simple multiplication in the scale domain (lower level). At each stage, one of a pair of *dual* filters is used, depending on whether the process takes place in the spatial (left) or spatial-frequency (right) domains.

V. MOVING THE ATTENTION POINT

In this section, we address the question of computing a foveated image or local spectrum at a location (x_0, y_0) other than the origin and propose an efficient implementation based on steerable filters, avoiding the computational burden of a variant filtering.

A foveated image with attention point x_0, y_0 can be computed by first displacing the image to position (x_0, y_0) at the origin and then applying the lowpass SRVF operator \mathcal{L}_P as in (2):

$$\begin{aligned} c_{x_0, y_0}(r, \theta) &= \mathcal{L}_P\{f(x + x_0, y + y_0)\} \\ &= \iint_{\mathcal{R}^2} dx dy f(x + x_0, y + y_0) \\ &\quad \cdot \frac{1}{r^2} P\left(\frac{x \cos \theta + y \sin \theta}{r}, \frac{y \cos \theta - x \sin \theta}{r}\right) \end{aligned} \quad (34)$$

where $c_{x_0, y_0}(r, \theta)$ is the foveated image around the attention point (x_0, y_0) . By defining

$$P_{r, \theta}(x, y) = \frac{1}{r^2} P\left(\frac{-x \cos \theta - y \sin \theta}{r}, \frac{x \sin \theta - y \cos \theta}{r}\right) \quad (35)$$

we can rewrite (34) as a convolution of $f(x, y)$ with scaling

and rotating *lowpass* filters $P_{r,\theta}(x, y)$:

$$\begin{aligned} c_{x_0, y_0}(r, \theta) &= \iint_{\mathcal{R}^2} dx dy f(x, y) P_{r,\theta}(x_0 - x, y_0 - y) \\ &= [f * P_{r,\theta}](x_0, y_0). \end{aligned} \quad (36)$$

Equivalently, following (28), the local spectrum estimated around (x_0, y_0) can be expressed as

$$C_{x_0, y_0}(\omega, \Omega) = [f * \tilde{p}_{\omega, \Omega}](x_0, y_0) \quad (37)$$

which again is a convolution between $f(x, y)$ and a continuum of scaling-rotating *bandpass* filters

$$\begin{aligned} \tilde{p}_{\omega, \Omega}(x, y) &= \tilde{p}(-\omega(x \cos \Omega + y \sin \Omega) \\ &\quad \omega(x \sin \Omega - y \cos \Omega)). \end{aligned} \quad (38)$$

A. Implementation

Equations (36) and (37) express a displaced fovea or its associated local spectrum as convolutions of the image $f(x, y)$ with either lowpass ($P_{r,\theta}$) or bandpass ($\tilde{p}_{\omega, \Omega}$) filters, respectively. However, despite the formal simplification, a direct implementation of those equations is still impractical since an infinite (or, in practice, very large) number of 2-D convolutions are needed to know either c_{x_0, y_0} or C_{x_0, y_0} at any position.

Nevertheless, the problem of an efficient implementation can be addressed through the use of steerable or deformable filters [22]–[24]. These techniques approximate, through a linear combination of a finite basis, a continuum of filters having any desired value in their parameters (scale and orientation, in our case). The coefficients in the linear combination (or interpolation functions) are obtained from the corresponding set of normal equations arising from the problem of minimizing the error (in a LSE sense) within the parameter range considered. In our case, the effective range of scales is determined in practice by the maximum resolution of the image and the spatial support of the fovea (note that, as opposed to the orientation case, a perfect scale interpolation is not possible due to the lack of periodicity of this variable [24], [25]).

In order to estimate the local spectrum, we follow the techniques described in [26] to obtain a basis of N (scale) $\times M$ (orientation) bandpass filters $\tilde{p}_{\omega_i, \Omega_j}(x, y)$ approximating any intermediate $\tilde{p}_{\omega, \Omega}$ as

$$\tilde{p}_{\omega, \Omega}(x, y) \simeq \sum_i^N \sum_j^M A_{ij}(\omega, \Omega) \tilde{p}_{\omega_i, \Omega_j}(x, y). \quad (39)$$

Then, the local spectrum $C_{x_0, y_0}(\omega, \Omega)$ is also approximated as a linear combination of $N \times M$ values at (x_0, y_0) from a set of bandpass filtered images $\text{BP}_{ij}(x, y)$ previously computed as invariant convolutions with the basis filters

$$\begin{aligned} C_{x_0, y_0}(\omega, \Omega) &= [f * \tilde{p}_{\omega, \Omega}](x_0, y_0) \\ &\simeq \sum_i^N \sum_j^M A_{ij}(\omega, \Omega) [f * \tilde{p}_{\omega_i, \Omega_j}](x_0, y_0) \\ &= \sum_i^N \sum_j^M A_{ij}(\omega, \Omega) \text{BP}_{ij}(x_0, y_0). \end{aligned} \quad (40)$$

The upper branch in Fig. 6 shows a local spectrum (central panel) around an attention point overlaid in the original image (left). It has been interpolated from only 8×8 samples in a log-polar grid (marked with \times 's) that constitute a complete (not exact but highly accurate) representation of the local spectrum. The filters $\tilde{p}_{\omega_i, \Omega_j}$ used in this example have radial and angular bandwidths of 1 octave and 40° , respectively, covering a range of about four octaves with an average signal-to-noise ratio of 30 dB (always above 25 dB). The foveated image can be obtained through an inverse Fourier transform of the local spectrum (Fig. 6 upper right). In practice, since the interpolated local spectrum is a linear combination, we have that

$$\begin{aligned} c_{x_0, y_0}(r, \theta) &= \mathcal{F}^{-1}\{C_{x_0, y_0}(\omega, \Omega)\} \\ &\simeq \sum_i^N \sum_j^M \text{BP}_{ij}(x_0, y_0) \mathcal{F}^{-1}\{A_{ij}(\omega, \Omega)\} \\ &= \sum_i^N \sum_j^M \text{BP}_{ij}(x_0, y_0) a_{ij}(r, \theta). \end{aligned} \quad (41)$$

Thus, the foveated image can be obtained directly as a linear combination of the inverse Fourier transform of the interpolation functions a_{ij} without explicitly computing the local spectrum.

However, if we just want to obtain a foveated image, a direct implementation in the spatial domain admits a further simplification by choosing a prototype filter having circular symmetry around its maximum ($x = 1, y = 0$). Then, there exists a one-variable function $R(\rho)$ that verifies

$$P(x, y) = P_{1,0}(-x, -y) = R((x-1)^2 + y^2). \quad (42)$$

Assuming circular symmetry (42), two filters having the same scale but with different orientations are translated versions of each other, that is, we can relate any $P_{r,\theta}$ to $P_{r,\theta=0}$ as

$$P_{r,\theta}(x, y) = P_{r,0}(x + r(1 - \cos \theta), y - r \sin \theta). \quad (43)$$

As a result, we do not have to compute the convolutions with the whole filterbank $P_{r,\theta}$ in (36) but only with the subset $P_{r,0}$. When the value of a foveated image $c_{x_0, y_0}(r, \theta)$ for an angle other than zero is needed, we look up $c_{x_\theta, y_\theta}(r, 0)$, which is the convolution with $P_{r,0}$, at the displaced position

$$x_\theta = x_0 + r(1 - \cos \theta), \quad y_\theta = y_0 - r \sin \theta. \quad (44)$$

Finally, we apply deformable-filter techniques but only to the scale parameter [27], [26], finding N filters $P_{r_i,0}$ that best approximate, in a LSE sense, any $P_{r,0}$ in the considered scale range

$$P_{r,0}(x, y) \simeq \sum_{i=1}^N b_i(r) P_{r_i,0}(x, y). \quad (45)$$

Therefore, the foveation algorithm in the spatial domain (lower branch in Fig. 6) also consists of two stages. First, we have the application of a filterbank $\{P_{r_i,0}\}$ to obtain N lowpass filtered images $\text{LP}_i(x, y) = [f * P_{r_i,0}](x, y)$. Since the convolution with any $P_{r,0}$ can be approximated using

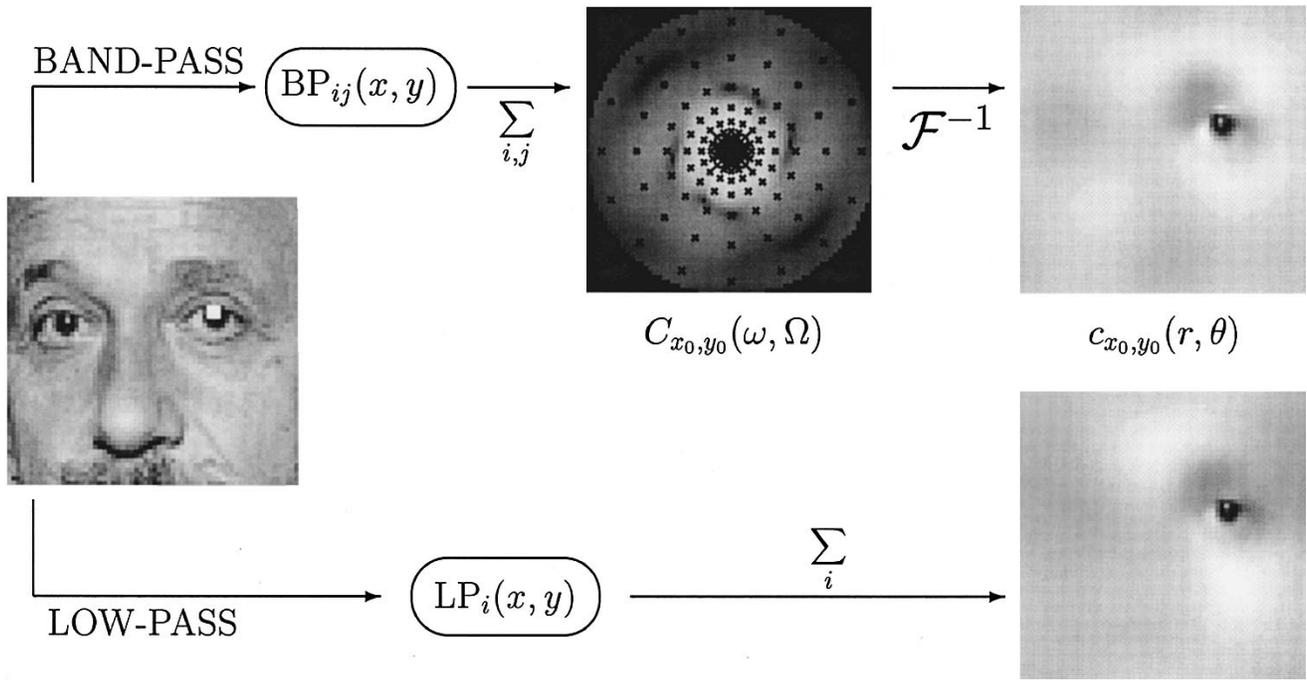


Fig. 6. Foveated images obtained in the spatial domain using scaled lowpass filters (lower branch) or through the local spectrum (upper branch) using scaled-rotated bandpass filters.

(45), then assuming circular symmetry, from (43) and (36), the displaced fovea is computed as

$$\begin{aligned}
 c_{x_0,y_0}(r, \theta) &= [f * P_{r,\theta}](x_0, y_0) = [f * P_{r,\sigma}](x_\theta, y_\theta) \\
 &\simeq \sum_i^N b_i(r) [f * P_{r_i,\sigma}](x_\theta, y_\theta) \\
 &= \sum_i^N b_i(r) LP_i(x_\theta, y_\theta). \tag{46}
 \end{aligned}$$

In the example of Fig. 6 (lower branch), we used only six Gaussian filters with circular symmetry, with a spatial support (measured at half height) of one octave, covering a range of five octaves with an average SNR of 25 dB. The resulting foveated images are very similar, the differences being due to the use of different kernels and to the wraparound effect in the frequency-domain implementation (upper panel). In terms of computational cost, the spatial implementation is significantly faster since fewer filterings, and hence fewer terms, are required in the linear combination (six versus 8×8 in the example of Fig. 6). With this spatial implementation, we have been able to move a fovea through a small static image in real time (several moves per second). This efficiency stems from the fact that once the initial lowpass filterings have been computed, moving the fovea only requires a linear combination of six terms at each point.

Fig. 7 displays two multifoveated images with different sets of attention points computed using the above algorithm. The original image (with the corresponding fixation points overlaid) is shown in the left panels. The right panels display the multifoveated images obtained by patching information from the different foveas. The original image is divided into patches (see left panels), and each area is substituted

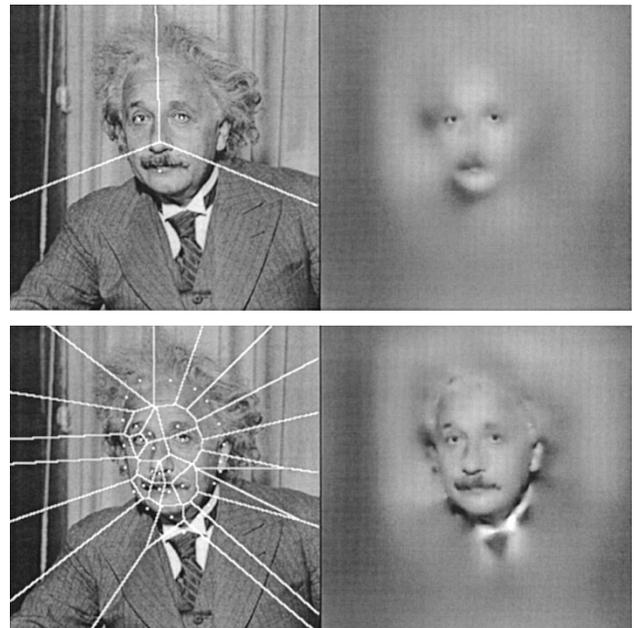


Fig. 7. Left: Original image showing the fixation points used (three in the upper panel and 30 in the lower one). Right: Reconstructed images using information from the nearest fovea.

by the foveated version obtained with the nearest fixation point. In the upper panel, only three points were considered, as opposed to the 30 foveas that were combined to show the result below. Fig. 7 shows how a relatively low number of attention points are enough for representing the most significant information of the picture. Considering the few samples required to represent each fovea (between 50 and 100, depending on the sampling scheme), the potential use of this method for dynamic compression [5] is promising, especially

in those applications where only the preservation of a few key features is necessary. Since each fovea is associated with a local spectrum, the images in the right panels of Fig. 7 can be seen as reconstructions from a set of local spectra estimated at different spatial locations.

VI. SUMMARY AND CONCLUSIONS

We have shown how a scaling-rotating variant filtering (SRVF) using a given prototype filter $P(r, \theta)$ in one domain is maintained as a similar SRVF in the conjugate domain but now using the “dual” filter \tilde{P} . This result was not evident, as it is very different to the well-known convolution theorem associated with an invariant filtering. Although the above discussion has been focused on images (2-D), the results are also valid in 1-D [18] and higher dimensions.

When the prototype filter is lowpass, this duality admits an interesting interpretation. A lowpass SRVF image can be straightforwardly realized as a “foveated” image—an important concept in vision—which involves the progressive degradation of the image as we move further away from the point of attention. On the other hand, the corresponding lowpass SRVF spectrum (or equivalently, the Fourier transform of the foveated image) corresponds to a local spectrum estimation around the point of fixation. Therefore, log-polar foveation and multiscale local spectral analysis can be regarded as similar processes, taking place in conjugated domains, and being related through a Fourier transform.

The estimation of either the foveated image or local spectrum has a much simpler formulation under a log-polar warping of both domains. Under this representation, the SRVF becomes an invariant filtering. Similarly, the visual system seems to perform a log-polar cortical mapping of the Cartesian object space while using log-distributed filters to sample the spectrum. Under this duality framework, an octave-based scheme to estimate the local spectrum naturally calls for a log-polar representation in the spatial domain [28]. Without intending it to be a model of the human visual system, we think that the idea that a multiscale estimation of the local spectrum mathematically implies a foveatization of the image, and vice-versa, could have application in the fields of vision research and artificial vision.

We also have presented a class of prototype filters generating self-dual kernels. Using these kernels, the SRVF is exactly the same in both domains. In other words, a SRVF operator using self-dual kernels commutes with the Fourier transform. It can be applied to any Fourier pair with the results being maintained as a Fourier pair. These self-dual filters are compatible with the requirements usually demanded of bandpass kernels for local spectrum estimation, including those used to model the receptive fields of simple neurons in the early stages of the visual cortex [12]. Furthermore, self-dual filters have the important feature of preserving simultaneously both the DC component (in the local spectrum) and the origin of the signal (in the foveated image). We have also presented a simple procedure to obtain a self-dual filtered signal from any well-behaved filter through a double application of the filter to both the signal and its spectrum.

The use of a log-polar warping also serves as an intermediate step to link our results with the scale transform. Based on this connection, we have proposed a generalization of Cohen’s 1-D scale transform to 2-D so that its modulus is invariant under scaling or rotation of the image (scale-orientation transform). Finally, we have presented an efficient implementation in both domains, based on steerable filters, that it is specially well suited for moving the attention point. With this implementation, it is possible to move a fovea across an image in real time.

There are still some aspects deserving further study, mainly those related to a theoretical formulation of the discrete case and the potential applications, especially those related to the study of rotation and scale invariances. Following our implementation, we obtain a discrete local representation combining spatial and spectral information that can be useful to define local feature descriptors, combining the invariances associated with the scale-rotation transform with the translation invariance of the local power spectrum. These descriptors could be used in tasks such as image segmentation, robust pattern recognition, etc. [26].

ACKNOWLEDGMENT

The authors would like to thank the editor and the anonymous reviewers whose comments led to a greatly improved manuscript.

REFERENCES

- [1] C. Braccini, G. Gambardella, and G. Sandini, “A signal theory approach to the space and frequency variant filtering performed by the human visual system,” *Signal Process.*, vol. 3, pp. 231–240, 1981.
- [2] C. A. Curcio, K. R. Sloan, R. E. Kalina, and A. E. Hendrikson, “Human photoreceptor topography,” *J. Comput. Neurol.*, vol. 292, pp. 497–523, 1990.
- [3] D. C. van Essen, W. T. Newsome, and J. H. R. Maunsell, “The visual field representation in striate cortex of the macaque monkey,” *Vision Res.*, vol. 24, pp. 429–448, 1984.
- [4] D. G. Green, “Regional variations in the visual acuity for interference fringes on the retina,” *J. Physiol.*, vol. 207, pp. 351–356, 1970.
- [5] P. Kortum and W. Geisler, “Implementation of a foveated image coding system for image bandwidth reduction,” *Proc. SPIE*, vol. 2657, pp. 350–360, 1996.
- [6] H. Yamamoto, Y. Yeshurun, and M. D. Levine, “An active foveated vision system: attention mechanisms and scan path convergence measures,” *Comput. Vision Image Understanding*, vol. 63, pp. 50–65, 1996.
- [7] W. Geisler and J. S. Perry, “A real-time foveated multiresolution system for low-bandwidth video communication,” *Proc. SPIE*, vol. 3299, pp. 294–305, 1998.
- [8] D. L. Jones and T. W. Parks, “A high-resolution data-adaptive time-frequency representation,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 2127–2135, 1990.
- [9] S. G. Mallat, “A theory for multiresolution signal decomposition: The wavelet transform,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, 1989.
- [10] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.
- [11] G. Cristobal and L. Cohen, “Scale in images,” *SPIE Proc. Adv. Signal Process.*, vol. 2846, pp. 251–261, 1996.
- [12] D. A. Pollen and S. F. Ronner, “Visual cortical neurons as localized spatial filters,” *IEEE Trans. Syst., Man, Cybern.*, vol. 13, pp. 907–916, 1983.
- [13] J. G. Daugman, “Spatial visual channels in the Fourier plane,” *Vision Res.*, vol. 24, pp. 891–910, 1984.
- [14] D. J. Field, “Relation between the statistics of natural images and the response properties of cortical cells,” *J. Opt. Soc. Amer. A*, vol. 4, pp. 2379–2394, 1987.
- [15] L. Cohen, “The scale representation,” *IEEE Trans. Signal Processing*, vol. 41, pp. 3275–3293, 1993.

- [16] W. Beil, "Steerable filters and invariance theory," *Patt. Recogn. Lett.*, 15, pp. 453–460, 1994.
- [17] H. Greenspan, S. Belongie, R. Goodman, P. Perona, S. Rakshit, and C. H. Anderson, "Overcomplete steerable pyramid filters and rotation invariance," in *Proc. IEEE Conf. Comput. Vis. Patt. Recogn.*, Seattle, WA, 1994, pp. 222–228.
- [18] A. Taberero, J. Portilla, and R. Navarro. "Duality between the local spectrum of a signal and its inverse Fourier transform, the local signal," Instituto de Óptica (CSIC), Tech. Rep. 53.
- [19] C. Gasquet and P. Witomski, *Analyse de Fourier et Applications*. Paris, France: Masson, 1990.
- [20] J. B. Allen and L. R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis," *Proc. IEEE*, vol. 65, pp. 1558–1564, 1977.
- [21] M. Michaelis and G. Sommer, "A Lie group-approach to steerable filters," *Pattern Recognit. Lett.*, vol. 16, pp. 1165–1174, 1995.
- [22] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 891–906, 1991.
- [23] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. Heeger, "Shiftable multiscale transforms," *IEEE Trans. Inform. Theory*, vol. 38, pp. 587–607, 1992.
- [24] P. Perona, "Deformable kernels for early vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 488–499, 1995.
- [25] P. C. Teo and Y. Hel-Or, "A computational group theory approach to steerable functions," Stanford Univ., Stanford, CA, Tech. Rep. STAN-CS-TN-96-33, Apr. 1996.
- [26] J. Portilla, R. Navarro, and A. Taberero, "Optimizing parametric deformable kernels," in *Proc. IEEE Int. Symp. Time-Scale Time-Freq. Anal.*, Pittsburgh, PA, Oct. 1998.
- [27] J. Portilla and R. Navarro, "Efficient method for space-variant lowpass filtering," in *Proc. SNRFAI*, Barcelona, Spain, 1997, pp. 287–292.
- [28] A. B. Watson, "Detection and recognition of simple spatial forms," in *Physical and Biological Processing of Images*, A. C. Sleight, Ed. Berlin, Germany: Springer-Verlag, 1982, pp. 100–114.



Antonio Taberero received the B.S. degree in physics in 1998 from the Universidad Complutense de Madrid, Madrid, Spain. He was a Research Assistant at the Instituto de Óptica, Universidad Politécnica de Madrid, from 1988 to 1992, when he received the Ph.D. degree from the Universidad Complutense de Madrid.

During 1993, he was a Postdoctoral Student at the Vision Group, Human Interface Research Branch, NASA Ames Research Center, Moffett Field, CA. He is currently a Member of the Faculty at the

Computer School, UPM. His research interests include computational vision, image processing, joint representations, and data compression.



early stages of human vision.

Javier Portilla received the B.S. degree in electrical engineering and computer science from Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad Politécnica de Madrid (UPM), Madrid, Spain, in 1994. He received the Ph.D. degree in 1999.

From 1995 to 1998, he was a Research Assistant with the Instituto de Óptica, UPM. His research interests include statistical models for visual texture, steerable filtering, and space-variant representations of images, which have always been inspired by the



is interested in human vision, optics, and image processing.

Rafael Navarro received the M.S. and Ph.D. degrees in physics from the University of Zaragoza, Zaragoza, Spain, in 1979 and 1984, respectively.

From 1985 to 1986, he was an Optical and Image Processing Engineer at the Instituto de Astrofísica de Canarias. He joined the Instituto de Óptica (CSIC), Universidad Politécnica de Madrid, in 1987, where he is currently a Senior Scientific Researcher. Since 1988, he has headed the Imaging and Vision Group, and since 1994, he has been Associate Director of the Instituto de Óptica. He