

ANISOTROPICALLY FOVEATED NONLOCAL IMAGE DENOISING

Alessandro Foi

Department of Signal Processing
Tampere University of Technology, Finland
<http://www.cs.tut.fi/~foi/>
alessandro.foi@tut.fi

Giacomo Boracchi

Dipartimento di Elettronica, Informazione
e Bioingegneria, Politecnico di Milano, Italy
<http://home.dei.polimi.it/boracchi/>
giacomo.boracchi@polimi.it

ABSTRACT

When our gaze fixates a point, the visual acuity is maximal at the fixation point (imaged by the fovea, i.e. the central part of the retina) and decreases rapidly towards the periphery of the visual field. This phenomenon is known as foveated vision or foveated imaging. We recently investigated the role of foveation in image filtering and we have shown that the foveated patch distance, i.e. the Euclidean distance between foveated patches, is a valuable feature for the assessment of nonlocal self-similarity. Foveation operators apply spatially variant blur, providing a compact multiscale representation of each image patch. Here, we introduce anisotropic foveation operators that embed directional point-spread functions, and we show that the operators providing the highest denoising quality are characterized by radial orientations. This result is coherent with the orientation preference in the human visual system.

Index Terms— Nonlocal Similarity, Image Denoising, Foveation, Human Visual System.

1. INTRODUCTION

Patch-based nonlocal imaging methods rely on the assumption that natural images contain a large number of mutually similar patches at different locations. Patch similarity is typically assessed through the Euclidean distance of the pixel intensities and therefore depends on the patch size: while large patches guarantee stability with respect to degradations such as noise, the mutual similarity that can be verified between pairs of patches tends to reduce as the patch size grows. Thus, a windowed Euclidean distance is commonly used to balance these two conflicting aspects, assigning lower weights to pixels far from the patch center.

In [1], patch foveation was proposed as an alternative to windowing in nonlocal imaging. Patch foveation is performed through a foveation operator, which consists in a spatially variant blur where the point-spread functions (PSFs) have bandwidth decreasing with the spatial distance from the patch center.

Patch similarity can be assessed by means of the foveated distance, i.e. the Euclidean distance of foveated patches, leading to the concept of foveated self-similarity. This approach is inspired by the human visual system (HVS), which features spatially variant properties [2, 3, 4, 5]: if we treat the center of the patch as a fixation point, the foveated distance mimics the inability of the HVS to perceive details at the periphery of the center of attention.

In contrast with the conventional windowing, which is only spatially selective and attenuates sharp details and smooth areas in equal way, patch foveation provides selectivity in both space and frequency domain. Foveated patches, in fact, embed pixels from fine-scale (the fixation point) to coarse-scale (pixels at the patch periphery) representations of patches, and such a space/frequency selectivity is achieved by means of non-adaptive operators.

In [1] we presented an explicit construction of a foveation operator yielding a foveated distance that, in terms of expectation under zero mean, i.i.d., Gaussian noise, is guaranteed to be equivalent to the distance induced by a given arbitrary windowing kernel. However, in presence of structured differences, such as those arising in the vicinity of edges, the windowed and foveated distances are fundamentally distinct, with the latter providing stronger response.

Therefore, the foveated self-similarity can be leveraged in a number of imaging applications and, in particular, the foveated patch distance was shown to be a valuable feature for assessing the nonlocal self-similarity in image filtering. The Foveated NL-means [1], which modifies the classical nonlocal means denoising filter (NL-means) [6] by computing the averaging weights based on the foveated patch distance instead of the conventional windowed patch distance, leads to a consistent improvement in the quality of the restored images according to both objective criteria and visual appearance, particularly due to better contrast and sharpness.

We here introduce an anisotropic generalization of the foveation operator, realized by means of directional PSFs. Fig. 1 compares an isotropic foveation operator based on circular Gaussian PSFs versus newly introduced anisotropic operators based on elliptical Gaussian PSFs. We show that the foveated distance induced by a specific class of anisotropic foveation operators provides further improvement in denois-

♥ This work was supported by the Academy of Finland (project no. 252547).

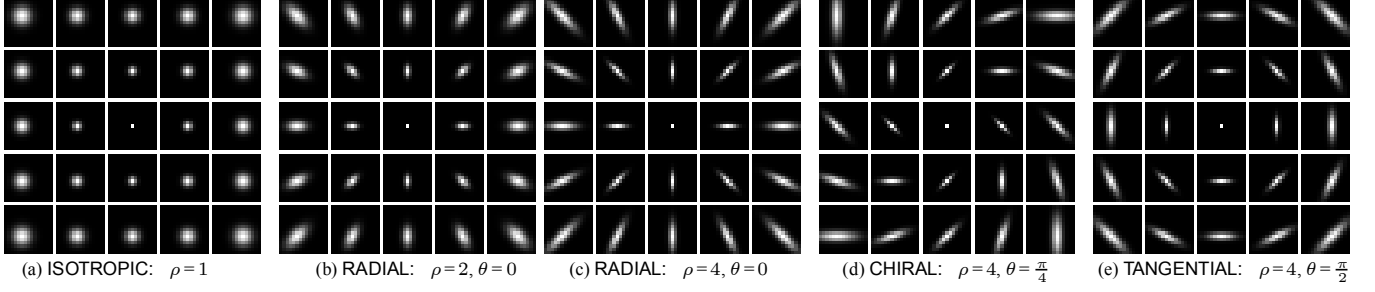


Fig. 1. Illustration of an isotropic foveation operator (a), and four anisotropic foveation operators (b - e). Each operator is displayed by the mosaic of the various PSFs that produce the pixels of a 5×5 foveated patch. The subimages in the mosaic are placed at the corresponding position of the pixels in the foveated patch. Note the relative displacement of the PSFs. For the sake of visualization, PSFs are displayed after intensity normalization.

ing performance with respect to the isotropic foveation operators. Remarkably, PSFs of anisotropic foveation operators that most successfully exploit the nonlocal similarity in natural images follow radial directions, in complete agreement with the orientation preference found in the HVS [7].

2. PRELIMINARIES

2.1. Observation Model

We consider noisy grayscale images $z : X \rightarrow \mathbb{R}$ modeled as

$$z(x) = y(x) + \eta(x), \quad x \in X \subset \mathbb{Z}^2, \quad (1)$$

where $X \subset \mathbb{Z}^2$ is a regular pixel grid (the image domain), $y : X \rightarrow \mathbb{R}$ is the unknown original noise-free image, and $\eta : X \rightarrow \mathbb{R}$ is i.i.d. Gaussian white noise, $\eta(\cdot) \sim \mathcal{N}(0, \sigma^2)$.

Noisy and noise-free patches centered at a pixel $x \in X$ are extracted from z or in y , respectively, as

$$\mathbf{z}_x(u) = z(u+x), \quad \mathbf{y}_x(u) = y(u+x), \quad u \in U,$$

where $U \subset \mathbb{Z}^2$ is a neighborhood of the origin.

2.2. Nonlocal Means (NL-means)

The purpose of any denoising algorithm is to provide an estimate \hat{y} of the original image y . In the basic implementation of NL-means [6] the denoised image \hat{y} consists in a weighted average of potentially¹ all the image pixels, i.e.

$$\hat{y}(x_1) = \sum_{x_2 \in X} w(x_1, x_2) z(x_2), \quad \forall x_1 \in X, \quad (2)$$

where $\{w(x_1, x_2)\}_{x_2 \in X}$ is the set of adaptive weights that characterize the pixel x_1 , which are positive and sum to one. Each weight $w(x_1, x_2)$ is determined by the similarity between the two patches \mathbf{z}_{x_1} and \mathbf{z}_{x_2} , as

$$w(x_1, x_2) = e^{-\frac{d(x_1, x_2)}{h^2}} / \sum_{x \in X} e^{-\frac{d(x_1, x)}{h^2}}, \quad (3)$$

¹In practice, mainly due to computational aspects, for each x_1 , only pixels x_2 belonging to a limited search window around x_1 are considered in (2).

where $d(x_1, x_2)$ is a distance measure between image patches centered at x_1 and x_2 , and $h > 0$ is a filtering parameter controlling the decay of the exponential function.

The distance d is defined as a windowed quadratic distance between patches, i.e.

$$d(x_1, x_2) = \left\| \mathbf{z}_{x_1} \sqrt{\mathbf{k}} - \mathbf{z}_{x_2} \sqrt{\mathbf{k}} \right\|_2^2 \quad (4)$$

being \mathbf{k} a non-negative windowing kernel defined over U , which adjust the contribution of each term depending on its position with respect to the patch center. Typically, \mathbf{k} is rotational symmetric and the weights $\mathbf{k}(u)$ are determined by the spatial distance from the center.

Equation (3) assigns larger weights to the terms $z(\cdot)$ in (2) that correspond to pixels belonging to similar patches (i.e. where the distance between patches $d(x_1, x_2)$ is small), regardless of their location within the image: thus, the NL-means enforces the self-similarity of natural images, which turns out to be an effective regularity prior for suppressing the noise.

2.3. Foveated distance

In the Foveated NL-means [1], the windowed patch distance $d(x_1, x_2)$ in (4) is replaced by the *foveated distance*

$$d^{\text{FOV}}(x_1, x_2) = \left\| \mathcal{F}[z, x_1] - \mathcal{F}[z, x_2] \right\|_2^2 = \left\| \mathbf{z}_{x_1}^{\text{FOV}} - \mathbf{z}_{x_2}^{\text{FOV}} \right\|_2^2, \quad (5)$$

where \mathcal{F} is a *foveation operator* that, given an image z and a fixation point x , outputs a *foveated patch* $\mathbf{z}_x^{\text{FOV}}$

$$\mathcal{F}[z, x](u) = \mathbf{z}_x^{\text{FOV}}(u), \quad u \in U. \quad (6)$$

In practice, \mathcal{F} corresponds to a spatially variant blurring operator with increasing blur (i.e. decreasing bandwidth) at pixels far from the center. Strictly speaking, $\mathbf{z}_x^{\text{FOV}}$ is, compared to \mathbf{z}_x , progressively blurrier as its argument $|u|$ grows, Fig. 1(a) shows the PSFs yielding the spatially variant blur of the foveation operators used in [1]. In the Foveated NL-means, the weights $\{w^{\text{FOV}}(x_1, x_2)\}$ are computed as in (3) replacing d^{FOV} with d . Fig. 2 shows an example of foveated patch and compares the weights w^{FOV} with w .

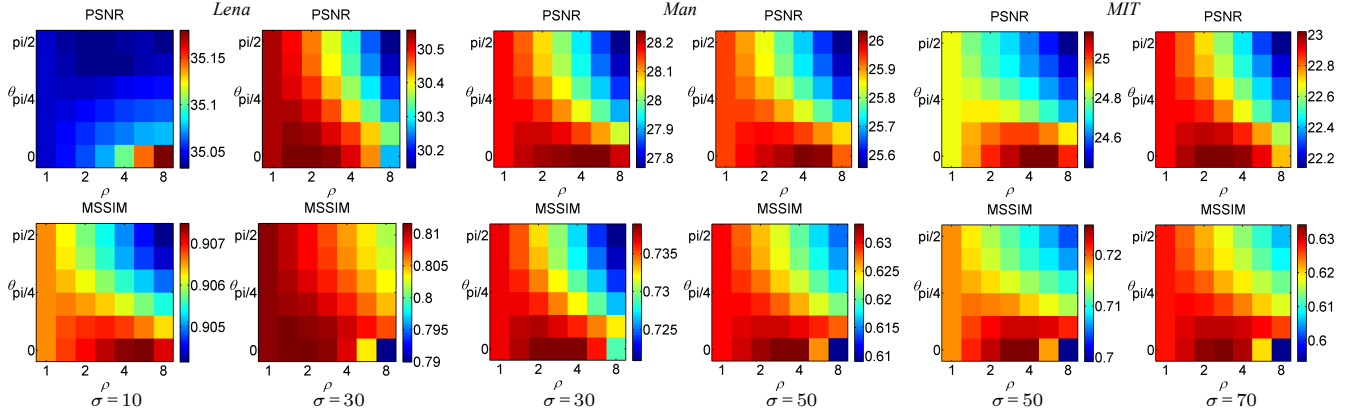


Fig. 3. Denoising performance in terms of PSNR and SSIM of the Foveated NL-means when the patch distance is measured by using anisotropic foveation operators. The best restoration performance can be achieved by anisotropic foveation operators characterized by PSFs elongated along the meridian lines toward the patch center (i.e. $\theta = 0$). In contrast, operators with tangential PSFs (i.e. $\theta = \pi/2$) are particularly ineffective and even lead to a performance loss with respect to the isotropic ones. As a figure of merit, the PSNR/SSIM results obtained by the NL-means (using windowed distance) for these six inputs are, respectively, 34.8/0.90, 29.2/0.79, 27.0/0.70, 24.9/0.60, 23.4/0.68, 20.9/0.58.

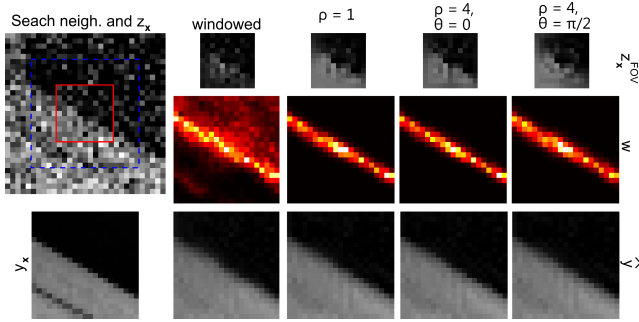


Fig. 2. First column: a portion of z (*MIT* image $\sigma=50$) highlighting the 21×21 search neighborhood at x (dashed blue line), and the 11×11 patch \mathbf{z}_x (solid red line) and, below, the original image in the search neighborhood. The remaining columns concern standard NL-means (windowed) and Foveated NL-means using *isotropic*, *radial*, and *tangential* foveation operators. The top row shows the patches used for computing weights, i.e., $\mathbf{z}_x \sqrt{\mathbf{k}}$ and $\mathbf{z}_x^{\text{FOV}}$, the middle and bottom rows illustrate the corresponding weights and the denoised estimates in the search neighborhood, respectively.

2.4. Foveation Operator Requirements

The foveation operator \mathcal{F} corresponding to a given windowing kernel \mathbf{k} can be constructed in such a way to satisfy the following four requirements [1].

Linearity. \mathcal{F} is linear with respect to the input image.

Flat-field preservation. \mathcal{F} maps a flat image into flat patches; i.e.

$$\exists \alpha > 0 \text{ such that } \forall c > 0 \text{ if } z(x) = c \quad \forall x \in X, \text{ then} \\ \mathcal{F}[z, x](u) = \alpha c \quad \forall u \in U, \quad \forall x \in X. \quad (7)$$

Central acuity. \mathcal{F} is fully sharp at the center of the patch:

$$\exists \alpha > 0 : \mathcal{F}[z, x](0) = \alpha z(x) \quad \forall x \in X.$$

Compatibility. The foveated distance d^{FOV} can be used as a straightforward replacement of the windowed distance d computed from \mathbf{k} in (3), without need of tuning parameters or other adjustments. Compatibility is defined for the comparison of perfectly identical patches, where we want d^{FOV} and d to coincide in mathematical expectation. Since $\mathbf{y}_{x_1} = \mathbf{y}_{x_2}$ implies $E\{d(x_1, x_2)\} = 2\sigma^2 \sum_{u \in U} \mathbf{k}(u)$, we formulate the compatibility requirement as follows: if $\mathbf{y}_{x_1}^{\text{FOV}} = \mathbf{y}_{x_2}^{\text{FOV}}$ then

$$E\{d^{\text{FOV}}(x_1, x_2)\} = E\left\{\|\mathbf{z}_{x_1}^{\text{FOV}} - \mathbf{z}_{x_2}^{\text{FOV}}\|_2^2\right\} = 2\sigma^2 \|\mathbf{k}\|_1. \quad (8)$$

In other words, (8) means that for any pair of identical noise-free foveated patches $\mathcal{F}[y, x_1] = \mathcal{F}[y, x_2]$, the expectation of the foveated distance $d^{\text{FOV}}(x_1, x_2)$, evaluated on the corresponding noisy foveated patches as in (5), should coincide with the expectation of the windowed distance (4) under comparable ideal conditions.

2.5. Construction of an Isotropic Foveation Operator

The construction of a foveation operator \mathcal{F} that satisfies the above requirements amounts to adjusting the scaling and spread of the blur PSFs in such a way that their ℓ^1 norm is constant and their squared ℓ^2 norms equal the corresponding values of the windowing kernel \mathbf{k} . For the particular case of circular-symmetric Gaussian PSFs (see Fig. 1(a)), the isotropic operator is defined by [1]

$$\mathcal{F}[z, x_1](u) = \sum_{x \in X} z(x) v_{\zeta(\mathbf{k}(u))}(x - x_1 - u), \quad \forall u \in U, \quad (9)$$

where

$$v_{\zeta(\mathbf{k}(u))}(x) = \sqrt{\mathbf{k}(0)} \tilde{v}_{\zeta(\mathbf{k}(u))} = \frac{2\mathbf{k}(u)}{\sqrt{\mathbf{k}(0)}} e^{-\frac{2\pi\mathbf{k}(u)|x|^2}{\mathbf{k}(0)}},$$

$\tilde{v}_{\zeta(\mathbf{k}(u))}$ being the bivariate Gaussian probability density function (p.d.f.) with mean zero and diagonal covariance matrix $\Sigma_{\zeta(\mathbf{k}(u))} = \mathbb{I}\zeta^2(\mathbf{k}(u))$, and $\zeta(\mathbf{k}(u)) = \sqrt{\frac{1}{4\pi} \frac{\mathbf{k}(0)}{\mathbf{k}(u)}}$. The ℓ^1 norm and the squared ℓ^2 norm of $v_{\zeta(\mathbf{k}(u))}$ are, respectively,

$$\|v_{\zeta(\mathbf{k}(u))}\|_1 = \alpha = \sqrt{\mathbf{k}(0)}, \quad \|v_{\zeta(\mathbf{k}(u))}\|_2^2 = \mathbf{k}(u). \quad (10)$$

Such foveation is said to be isotropic because all blur PSFs are circular symmetric and attenuate image features regardless of their orientation.

3. ANISOTROPIC FOVEATION OPERATORS

Anisotropic foveation operators generalize the isotropic ones described in Section 2.5 by utilizing, in place of the circular-symmetric Gaussian p.d.f. $\tilde{v}_{\zeta(\mathbf{k}(u))}$, an elliptical Gaussian p.d.f. whose covariance matrix depends not only on $\zeta(\mathbf{k}(u))$, but also on u itself, on a parameter $\rho > 0$ which determines the elongation of the p.d.f., and on an angular-offset parameter θ which controls the orientation of the axes of the elliptical p.d.f. Specifically, we denote by $\tilde{v}_{\zeta(\mathbf{k}(u)),\rho,\theta,u}$ the elliptical Gaussian p.d.f. whose covariance matrix has the form

$$\Sigma_{\zeta(\mathbf{k}(u)),\rho,\theta,u} = \zeta^2(\mathbf{k}(u)) R_{\angle u+\theta} D_{\rho} R_{\angle u+\theta}^T,$$

where the matrix $D_{\rho} = \begin{bmatrix} \rho & 0 \\ 0 & 1/\rho \end{bmatrix}$ determines the elongation and $R_{\angle u+\theta} = \begin{bmatrix} \cos(\angle u+\theta) & -\sin(\angle u+\theta) \\ \sin(\angle u+\theta) & \cos(\angle u+\theta) \end{bmatrix}$ is a rotation matrix of angle $\angle u + \theta$; in the special case $u = 0$, for which $\angle u$ cannot be defined, we impose $\rho = 1$, thus reducing to a circular-symmetric PSF. Overall, the foveation operator (9) becomes

$$\mathcal{F}_{\rho,\theta}[z, x_1](u) = \sum_{x \in X} z(x) \sqrt{\mathbf{k}(0)} \tilde{v}_{\zeta(\mathbf{k}(u)),\rho,\theta,u}(x - x_1 - u).$$

When $\theta = 0$ and $\rho > 1$, the major axes of the PSFs are directed along the meridian lines toward the patch center, yielding *radial* foveation operators. Conversely, if $\theta = \pi/2$ and $\rho > 1$, it is the minor axes of the PSFs that are directed towards the patch center and we obtain *tangential* foveation operators. Due to symmetry, the operators $\mathcal{F}_{\rho,\theta}$, $\mathcal{F}_{1/\rho,\theta+\pi/2+n\pi}$ and $\mathcal{F}_{\rho,\theta+n\pi}$ coincide for any $n \in \mathbb{Z}$. The larger $\max\{\rho, 1/\rho\}$ is, the more elongated the PSFs are along their major axis; when $\rho = 1$, $\mathcal{F}_{\rho,\theta}$ coincides with the *isotropic* foveation operator (9). Fig. 1 shows the PSFs that constitute a few anisotropic foveation operators corresponding to the same \mathbf{k} for different values of ρ and θ .

3.1. The Four Constraints

The anisotropic foveation operators $\mathcal{F}_{\rho,\theta}$ satisfy the four constraints set in Section 2.4 for any combination $\rho > 0$ and $\theta \in (-\pi/2, \pi/2]$. The proof, which we omit because of length limitation, is rather straightforward and boils down to the fact that the norms (10) are invariant with respect to elliptical deformation of the probability density functions $\tilde{v}_{\zeta(\mathbf{k}(u))}$ as long as the covariance matrix has always unitary determinant, which is ensured by the definition of D_{ρ} .

4. EXPERIMENTS

We assess the effectiveness of anisotropic foveation by denoising natural test images corrupted by additive Gaussian noise using Foveated NL-means. In particular, we consider operators $\mathcal{F}_{\rho,\theta}$ with various elongation and orientation combinations, varying $\rho \in [1, 8]$ and $\theta \in [0, \pi/2]$. Due to the low chirality of natural images, operators $\mathcal{F}_{\rho,\theta}$ and $\mathcal{F}_{\rho,-\theta}$ yield practically identical results, even for $\theta \in (0, \pi/2)$. Denoising performance for different levels of noise standard deviation σ are measured in terms of both PSNR and MSSIM [8]. Fig. 3 reports a representative portion of the obtained results.

Both PSNR and MSSIM results show that radial anisotropic patch foveation (i.e. $\theta = 0$, $\rho > 1$) improves the denoising performance with respect to the isotropic foveation ($\rho = 1$) or the anisotropic tangential foveation ($\theta = \pi/2$, $\rho > 1$). It emerges that radial foveation operators are more effective at assessing the nonlocal similarity in natural images. In particular, the performance gap between radial and isotropic foveation is larger on images characterized by marked edges, such as *MIT*.

The performance gap between radial and isotropic foveation is less substantial than the improvement achievable when introducing isotropic foveation in the classical, windowing-based, NL-means. Nevertheless, such improvement is particularly meaningful as it highlights that the radial foveation yields a stronger prior for measuring nonlocal similarity in image denoising. This result can be justified through arguments similar to those leading to the anisotropic NL-means [9], where it is shown that the by orienting and elongating the patch U along image edges it is possible to improve the noise suppression. As shown in Fig. 2, radial foveation preserves the substantial edge structure, since pixels are blurred along the edge rather than across the edge. The corresponding weights are in fact more precisely located along the edge direction. Such edge preservation is coherent with the results from [9]; in contrast, tangential foveation operators blur the patch across edges and result in a performance loss for all the considered images.

5. CONCLUSIONS

In this work we introduce anisotropic foveation operators and we show that assessing patch similarity using radial foveation is a more effective regularity assumption than using the isotropic foveation or the windowing conventionally used in NL-means. In terms of computational complexity, foveated NL-means introduces a negligible overhead, since foveated patches can be preliminary computed from the convolution of z against the PSFs $\tilde{v}_{\zeta(\mathbf{k}(u)),\rho,\theta,u}(x)$ of the corresponding foveation operator.

More interestingly, this radial PSF layout agrees with the orientation preference found at various levels of the human visual system [7],[10].

6. REFERENCES

- [1] A. Foi and G. Boracchi, "Foveated self-similarity in nonlocal image filtering," in *SPIE Human Vision and Electronic Imaging XVII*, vol. 8291. Burlingame, California, USA: SPIE, 2012.
- [2] C. Curcio, K. Sloan, R. Kalina, and A. Hendrickson, "Human photoreceptor topography," *J. Comparative Neurology*, vol. 292, pp. 497–523, 1990.
- [3] T. Wertheim, "Über die indirekte sehschärfe," *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, vol. 7, pp. 172–187, 1894.
- [4] B. A. Wandell, *Foundations of Vision*. Sinauer Associates, Inc., 1995.
- [5] M. P. Eckstein, "Visual search: A retrospective," *Journal of Vision*, vol. 11, no. 5, 2011. [Online]. Available: <http://www.journalofvision.org/content/11/5/14.abstract>
- [6] A. Buades, B. Coll, and J. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling Simulation*, vol. 4, no. 2, p. 490, 2005.
- [7] T. Sasaki, R. Rajimehr, B. Kim, L. Ekstrom, W. Vanduffel, and R. Tootell, "The radial bias: a different slant on visual orientation sensitivity in human and nonhuman primates," *Neuron*, vol. 51, no. 5, pp. 661–670, September 2006.
- [8] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, April 2004.
- [9] A. Maleki, M. Narayan, and R. G. Baraniuk, "Anisotropic nonlocal means," *Applied and Computational Harmonic Analysis*, In Press. [Online]. Available: <http://arxiv.org/abs/1112.0311>
- [10] J. Freeman, G. J. J. Brouwer, D. J. Heeger, and E. P. Merriam, "Orientation decoding depends on maps, not columns." *The Journal of Neuroscience*, vol. 31, no. 13, pp. 4792–4804, Mar. 2011.