Intelligent Embedded Systems

Advanced Research

# Foveated Self-Similarity in Nonlocal Image Filtering

Giacomo Boracchi

Dipartimento Elettronica, Informazione e Bioingegneria, Politecnico di Milano

giacomo.boracchi@polimi.it

Università Svizzera Italiana, 24 September 2013

Joint work with Alessandro Foi from Tampere University of Technology

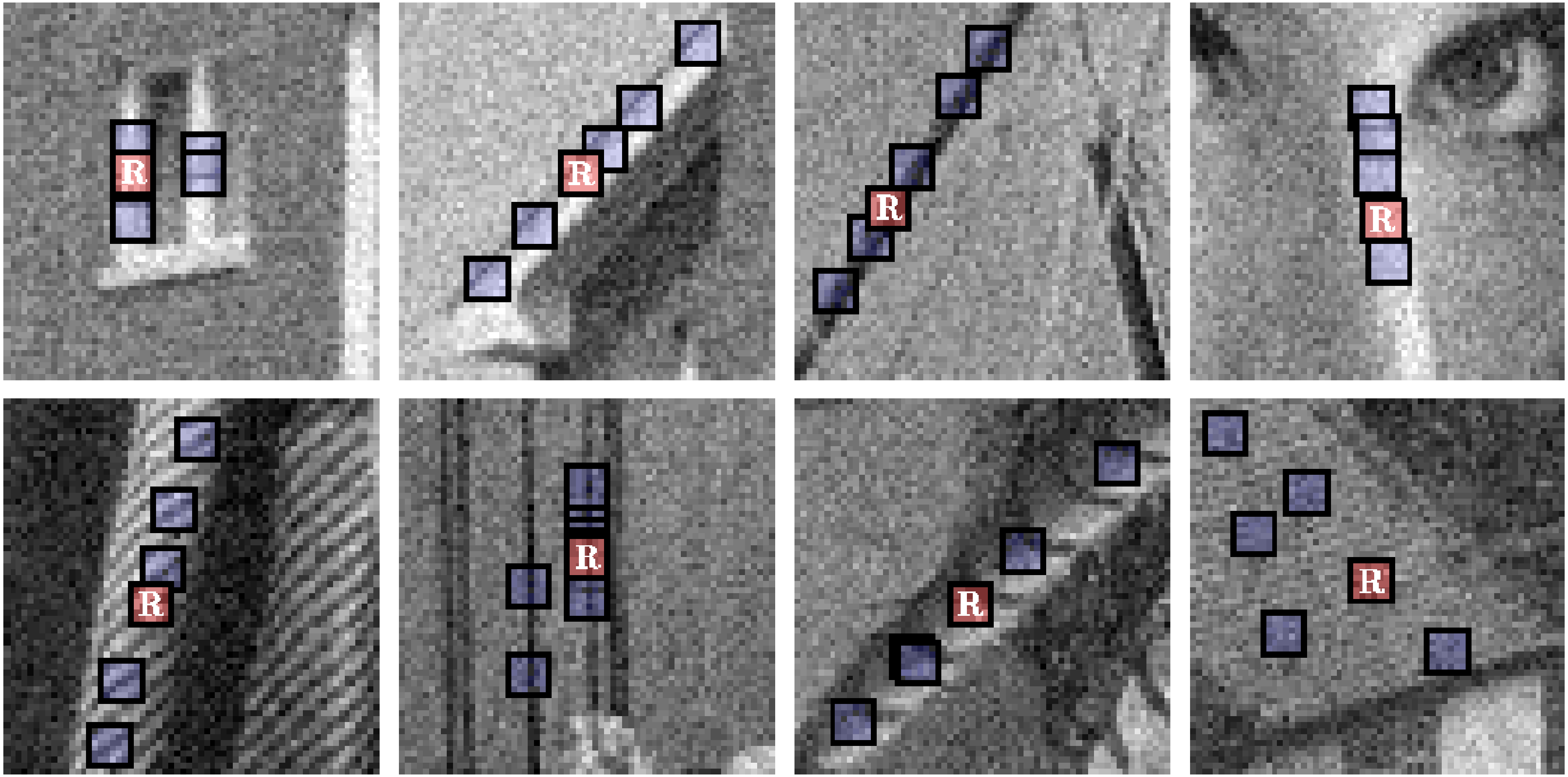
# Outline

- Nonlocal self similarity and Image Denoising
- Foveation and the Human Visual System
- Foveated Nonlocal Self Similarity
- Foveated NL-Means
- Experiments and Discussion
- Anisotropic Foveation
- Experiments and Discussion
- Conclusions

# Nonlocal Self Similarity

In Image processing, a brief introduction

# NonLocal Self Similarity




*In a natural image, for any given patch there exist **many** other **similar** looking **patches** at **different spatial locations**.*

# NonLocal Self Similarity in Image Processing

- Traced back to **fractal models** of natural images (Barnsley, 1993) and fractal block coding (Jacquin, 1992)
  .. *self-transformability on a blockwise basis…*
- **Texture synthesis** and **completion** (Efros and Leung, 1999; Wei and Levoy, 2000).
- Predicting the **central pixel** of a patch by exploiting the *long-range correlation* of natural images (Zhang and Wang, 2002)
- Nonlocal self-similarity as **an effective regularity assumption** at the heart of many successful image **denoising algorithms** (NL-means, BM3D, etc.).
- Nonlocal self-similarity was successfully used for **several image/video processing tasks**.

# Image denoising (NL-Means)

a tool to quantitatively assess the performance of a descriptive model

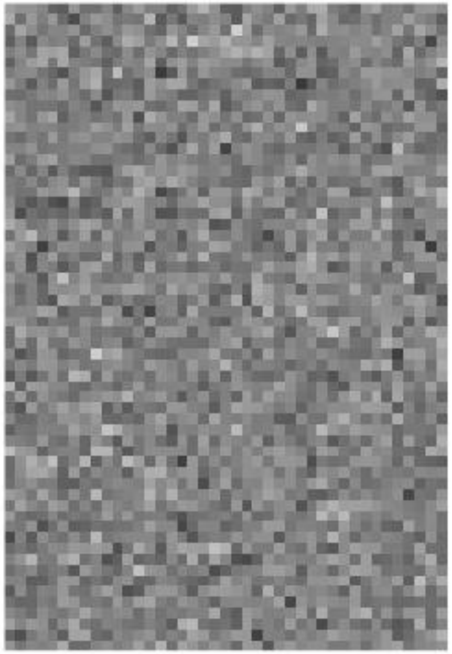
# Observation Model

$$z(x) = y(x) + \eta(x)\,, x \in X$$

- $z : X \to \mathbb{R}$ observed noisy image
- $y: X \to \mathbb{R}$ unknown original image (grayscale)
- $\eta: X \to \mathbb{R}$ i.i.d. Gaussian white noise, $\eta \sim N(0, \sigma^2)$

$z$

$y$



$\eta$

# Goal of image denoising

- The purpose of any **denoising** algorithm is to provide $\hat{y}$, an estimate of the original image $y$.

$z$

$\hat{y}$

# Goal of image denoising

- The purpose of any **denoising** algorithm is to provide $\hat{y}$, an estimate of the original image $y$.
- Denoising is an **ill posed problem** and requires some form of **regularization**.
- We consider **nonlocal self similarity** of image **patches**
- **Similar patches** have to be correctly identified on the basis of a suitable patch **distance measure**
- Such a distance implies the assumption of a **specific descriptive model for natural images** and their self-similarity.
- The denoising effectiveness actually depends on the validity of such underlying model.

# Patches

- Let $U \subset \mathbb{Z}^2$ be a spatial neighborhood centered at the origin $(0,0) \in \mathbb{Z}^2$,

  we define a patch centered at a pixel $x \in X$ in the observation $z$

$$\boldsymbol{z_x}(u) = z(x+u), \qquad u \in U$$

  a patch centered at a pixel $x \in X$ in the original image $y$

$$\boldsymbol{y_x}(u) = y(x+u), \qquad u \in U$$

# Non Local Means Filter (NL-means)

The denoised image $\hat{y}$ is a weighted average of all image pixels

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

where weights $\{w(x_1, x_2)\}$ are adaptively defined depending on the similarity between two noisy patches $\boldsymbol{z}_{x_1}$ and $\boldsymbol{z}_{x_2}$

$$w(x_1, x_2) = \frac{e^{\left(-\frac{d(x_1,x_2)}{h^2}\right)}}{\sum_X e^{\left(-\frac{d(x_1,x_2)}{h^2}\right)}}$$

- $d(x_1, x_2)$: distance measure between patches in $x_1$ and $x_2$,
- $h > 0$ is a smoothing parameter ($h = \sigma$).
- $\boldsymbol{z}_{x_1}$ similar to $\boldsymbol{z}_{x_2} \Rightarrow d(x_1, x_2)$ is small $\Rightarrow w(x_1, x_2)$ large
- NL-means operates pixel-wise

**A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," Multisc. Model. Simulat., vol. 4, no. 2, pp. 490-530, 2005**

# Windowed Patch Distance in NL-means

- The distance operator is defined as a *windowed quadratic distance* between patches

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

$$= \left\| \left(\boldsymbol{z}_{x_1} - \boldsymbol{z}_{x_2}\right)^2 \boldsymbol{k} \right\|_1 =$$

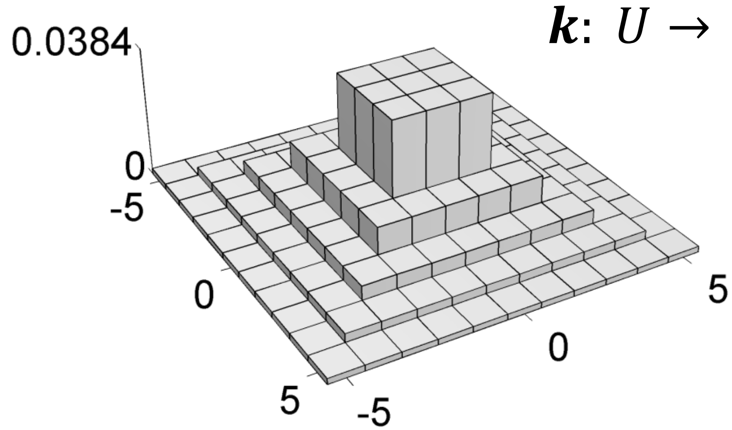
$$= \sum_{u \in U} \left( z(x_1 + u) - z(x_2 + u) \right)^2 \boldsymbol{k}(u)$$

- $\boldsymbol{k}$: $U \rightarrow \mathbb{R}^+$ is a windowing kernel

- The idea is to assess the similarity between pixels $y(x1)$ and $y(x2)$ (not available), through the similarity of the corresponding noisy patches $\boldsymbol{z}_{x_1}$ and $\boldsymbol{z}_{x_2}$.

# Windowed patch distance in NL-means (cnt.)

- The windowing kernel $\boldsymbol{k}\colon U \to \mathbb{R}^{+}$ adjusts the contribution of each difference term depending on the position of $u$ with respect to the patch center.




- $d$ performs a **pixel-wise** comparison of the patches
- the decay of $\boldsymbol{k}$ reflects how much similarity between $y(x_1)$ and $y(x_2)$ may be implied from the similarity between $y(x_1+u)$ and $y(x_2+u)$ when $u \neq 0$.

# Foveation

and The Human Visual System

# Lena foveated at two different fixation points

# Foveation in Image Processing

- **Image compression** (Kortum and Geisler, 1996): Any user gazing a screen would not notice significant differences between:
  - the fully detailed image properly displayed
  - the image foveated with respect to the fixation point.
- **Video compression** where fixation point can be tracked or estimated (Geisler and Perry, 1998; Lee et al., 2001; Basu and Wiebe, 1998).
- Image **coding** (Wang and Bovik, 2001) and video coding (Wang and Bovik, 2006).
- **Keypoint descriptor** (Alahi et al, 2012) inspider to the retina layout

# Foveated Nonlocal Self Similarity

and Foveation operators

# Foveated Self-Similarity

- **IDEA:** Replace windowing by foveation
- The windowed distance

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

- Is replaced by the **foveated distance**

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

where $\mathcal{F}$ is the **foveation operator** that, given an image $z$ and a fixation point $x$, outputs a foveated patch $\boldsymbol{z}_{x_1}^{\mathrm{FOV}}: U \to \mathbb{R}$, i.e.

$$\mathcal{F}[z, x_1](u) = \boldsymbol{z}_{x_1}^{\mathrm{FOV}}(u), u \in U$$

- **Foveation operators** reproduces **foveation effects** on image **patches** when the **fixation point** is the **patch centers**

# Foveation Operators

- Formally, $\mathcal{F}$ is a **space-variant blurring operator** with increasing blur (decreasing bandwidth) as we leave the center
- $\boldsymbol{z}_{x_1}^{\mathrm{FOV}}(u)$ is, compared to $\boldsymbol{z}_{x_1}$, progressively blurrier as $|u|$ grows

# Foveation Operators

- Formally, $\mathcal{F}$ is a **space-variant blurring operator** with increasing blur (decreasing bandwidth) as we leave the center
- $\boldsymbol{z}_{x_1}^{\mathrm{FOV}}(u)$ is, compared to $\boldsymbol{z}_{x_1}$, progressively blurrier as $|u|$ grows
- If we consider the **patch center as a fixation point**, $d^{\mathrm{FOV}} = \left\| \boldsymbol{z}_{\mathrm{x}_1}^{\mathrm{FOV}} - \boldsymbol{z}_{x_2}^{\mathrm{FOV}} \right\|_2^2$ mimics the **inability of the HVS to perceive details** at the **periphery** of the center of attention
- Foveation operators have to correspond to a specific windowing kernels.
- Thus, it is possible and easy to replace $d$ with $d^{\mathrm{FOV}}$

# Constrained Design of Foveation Operators

1. **Linearity and Translation Invariance:** $\mathcal{F}$ is a linear operator with respect to the image

$$\mathcal{F}[\lambda_1 z_1 + \lambda_2 z_2, x - \tau] = \lambda_1 \mathcal{F}[z(\cdot + \tau), x] + \lambda_2 \mathcal{F}[z(\cdot + \tau), x]$$

2. **Non-Negativity:** Foveated patches from non-negative images are non-negative

$$\text{if } z(x) > 0 \ \forall x \in X, \text{then } \mathcal{F}[z, x](u) \geq 0 \ \forall u \in U, \forall x \in X$$

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

$$\exists \alpha > 0 : \ \mathcal{F}[z, x](0) = \alpha z(x)$$

This property aims at mimicking the peak of the visual acuity at the fovea.

# Constrained Design of Foveation Operators (cnt)

4. **Flat-field preservation** $\mathcal{F}$ maps a flat image into a flat patches

$$\exists \alpha > 0 : \forall c > 0 \text{ if } z(x) = c \;\forall x \in X$$

$$\text{then } \mathcal{F}[z,x](u) = \alpha c \;\forall u \in U \;\forall x \in X$$

5. **Compatibility** $d^{\text{FOV}}$ can replace $d$ in NL-means, yielding the same expected distance in the **ideal case** where **perfectly identical patches** are compared.

*The mathematical expectation of the windowed distance operator is:* $E\{d(x_1,x_2)\} = E\left\{\left\|\left(\boldsymbol{z}_{x_1} - \boldsymbol{z}_{x_2}\right)^2 \boldsymbol{k}\right\|_1\right\} =$

$$\left\|\left(\boldsymbol{y}_{x_1} - \boldsymbol{y}_{x_2}\right)^2\right\|_1 + 2\sigma^2\|\boldsymbol{k}\|_1$$

## Constrained Design of Foveation Operators (cnt)

4. **Flat-field preservation** $\mathcal{F}$ maps a flat image into a flat patches

$$\exists\alpha > 0 : \forall c > 0 \text{ if } z(x) = c\ \forall x \in X$$

$$\text{then } \mathcal{F}[z,x](u) = \alpha c\ \forall u \in U\ \forall x \in X$$

5. **Compatibility** $d^{\mathrm{FOV}}$ can replace $d$ in NL-means, yielding the same expected distance in the **ideal case** where **perfectly identical patches** are compared.

$$\text{If } \boldsymbol{y}_{x_1}^{\mathrm{FOV}} = \boldsymbol{y}_{x_2}^{\mathrm{FOV}} \text{ then } E\left\{d^{\mathrm{FOV}}(x_1, x_2)\right\} = 2\sigma^2 \|\boldsymbol{k}\|_1$$

where $\boldsymbol{y}_x^{\mathrm{FOV}}$ denotes the noise-free foveated patches, i.e. $\mathcal{F}[y\ x](u) = \boldsymbol{y}_x^{\mathrm{FOV}}(u)\ \ \forall u \in U$

# Construction Of The Foveation Operator

- To satisfy linearity and non-negativity $\mathcal{F}$ admits the following representation

$$\mathcal{F}[y\ x](u) = \sum_{x \in X} z(x) v_u(x\ - x_1 - u), \qquad u \in U$$

i.e., is a linear blur translation-invariant w.r.t. $x_1$ and space-variant w.r.t $u$.

- The foveation operator is univocally determined by $\{v_u\}_{u \in U}$
- Thus $v_u > 0$ is a point-spread function (PSF) responsible for the blurring in the foveated patch at the position $u$.
- The standard-deviation (i.e. the spread) of $v_u$ is determined by the windowing kernel $\boldsymbol{k}$ in such a way to fulll the above four requirements.

# Construction of the foveation operator

- **IDEA:** construct, through suitable dilation and scaling, a family of kernels parametrized by $u$, where the kernels have same $\ell^1$ norm but varying $\ell^2$ norm determined by $\boldsymbol{k}(u)$.

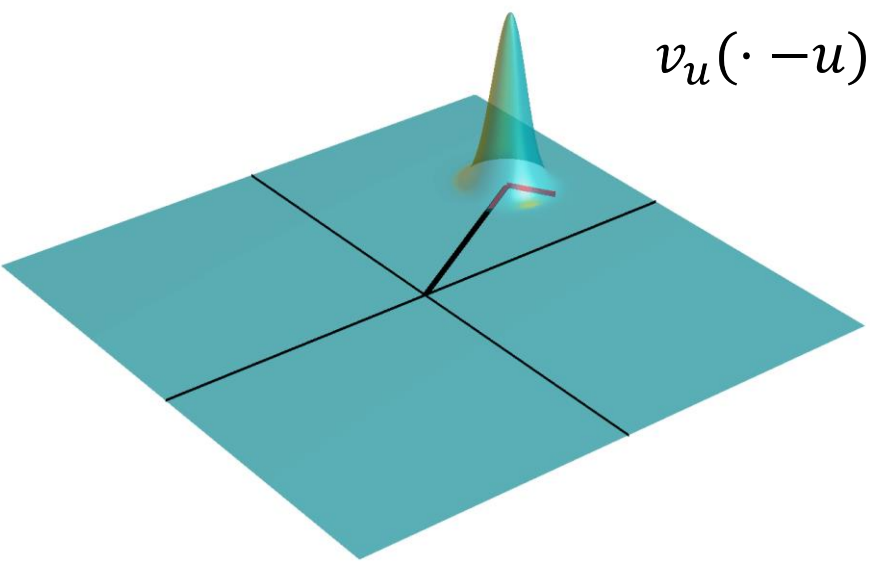
$$\left.\begin{array}{c} v_0 \approx \delta_0 \rightsquigarrow \text{central acuity} \\ \ell^1 \text{ norm} \equiv \alpha \rightsquigarrow \text{flat field preservation} \\ \ell^2 \text{ norm} = \sqrt{\boldsymbol{k}(u)} \rightsquigarrow \text{compatibility} \end{array}\right\} \alpha = \sqrt{\boldsymbol{k}(0)} \rightsquigarrow \text{compatibility}$$

- In what follows we consider foveation operators induced by Gaussian, circular symmetric PSFs

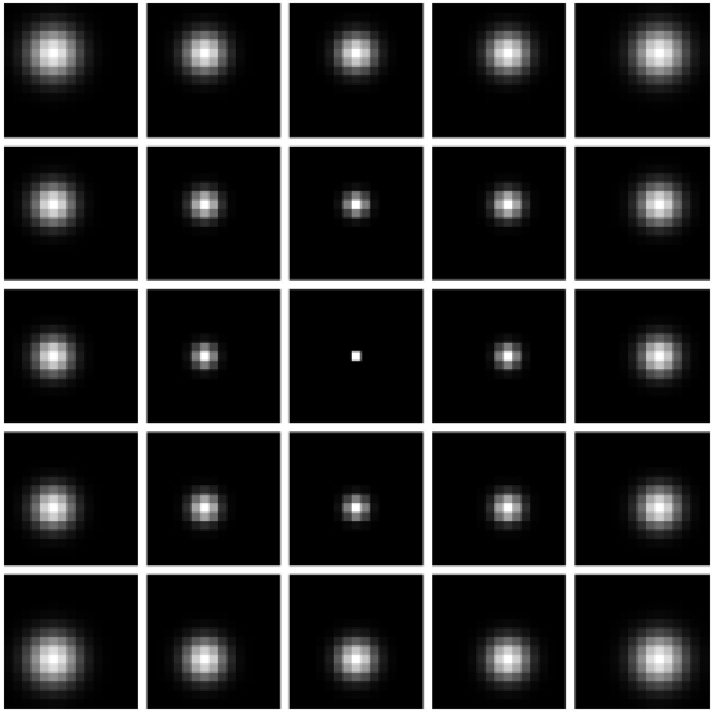
# Visualization of the foveation operator $\mathcal{F}$

- for a $5 \times 5$ windowing kernel $\boldsymbol{k}$

$$\mathcal{F}[y\,x](u) = \sum_{x \in X} z(x) v_u(x - x_1 - u), \qquad u \in U$$



$v_u(\cdot - u)$



# Foveated NL-means

a simple modification of NL-means

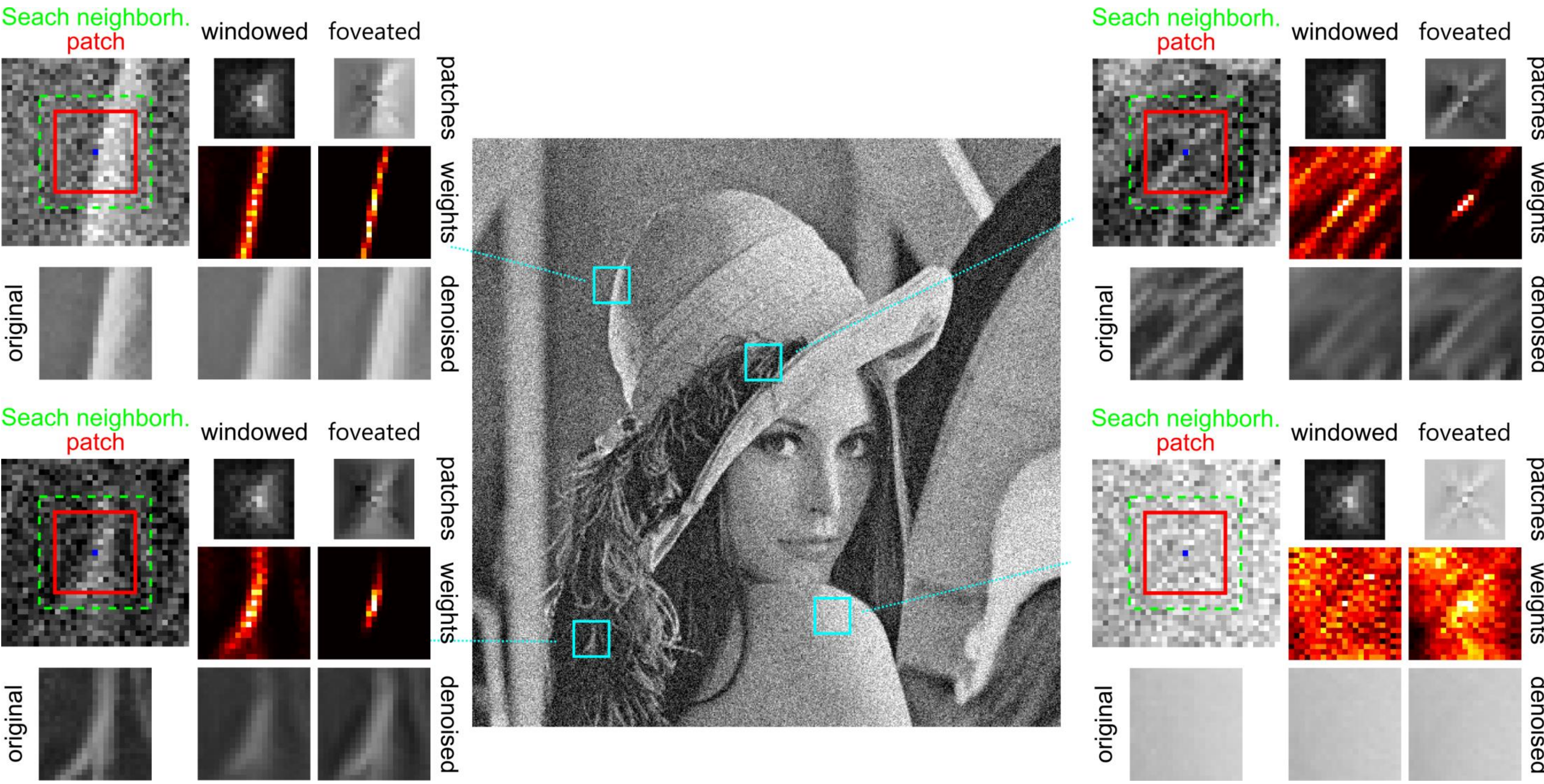
## Disclaimer:

- Our **goal** is **not** to introduce a **new denoising algorithm**.
- The removal of additive white Gaussian noise is the most widely used task for **quantitatively assessing the validity of any descriptive** or generative **model** of natural images.
- The **denoising performance** are here considered as a **compact indicator** of the **ability to identify similar patches** and to **distinguish** between **different ones** in noisy environments.

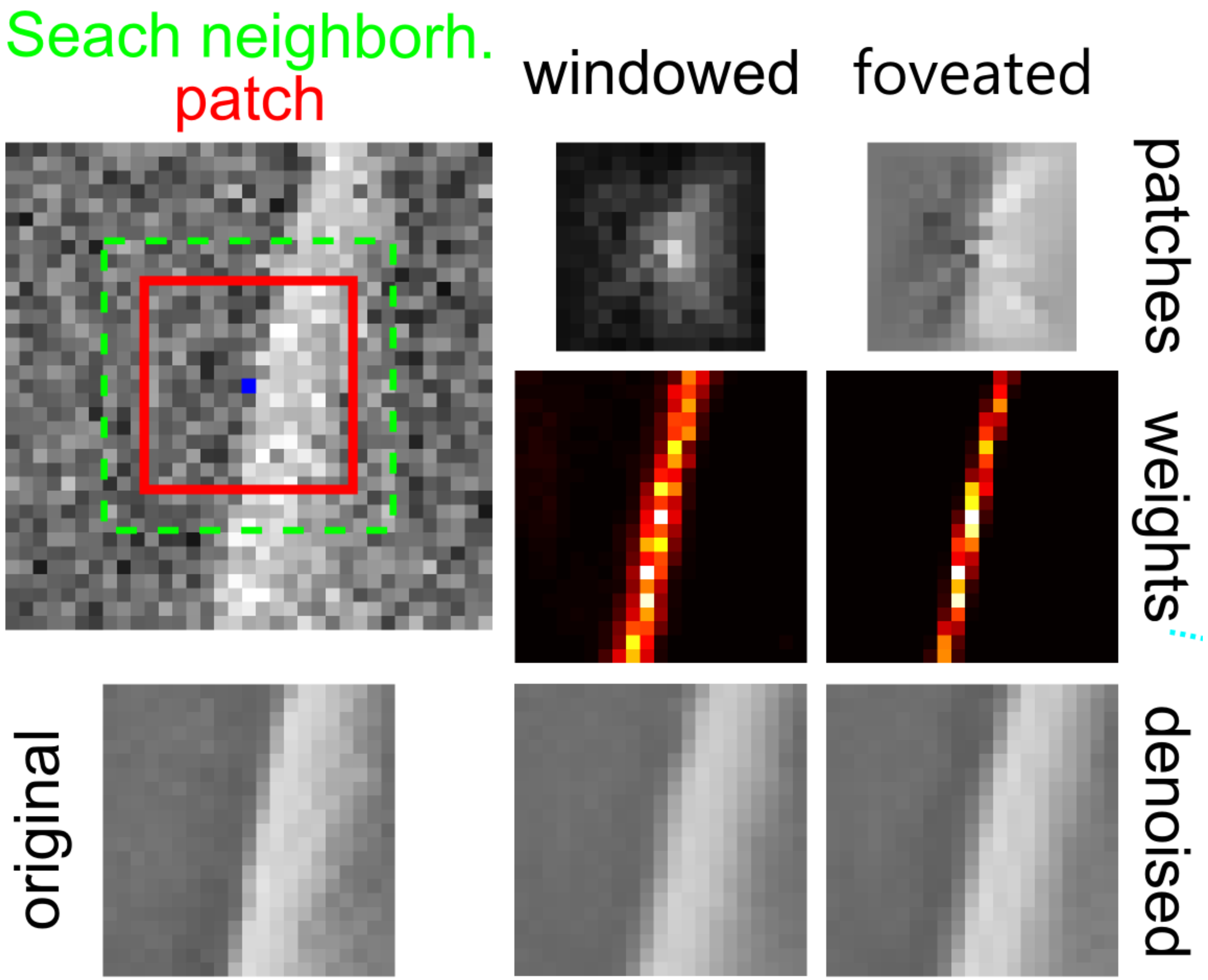
# Foveated NL-means

- Foveated NL-means is obtained from NL-means by **replacing** the **windowed distance** $d$ with the **foveated distance** $d^{\mathrm{FOV}}$ defined from them same windowing kernel $\boldsymbol{k}$
- **Compatibility** constraint ensures that the **two filters perform similarly** in areas where nearly **all patches are almost identical** to each other (ideal case of nonlocal self-similarity),**elsewhere the two filters depart from each other**
- **Matlab software at http://www.cs.tut.fi/~foi/FoveatedNL**

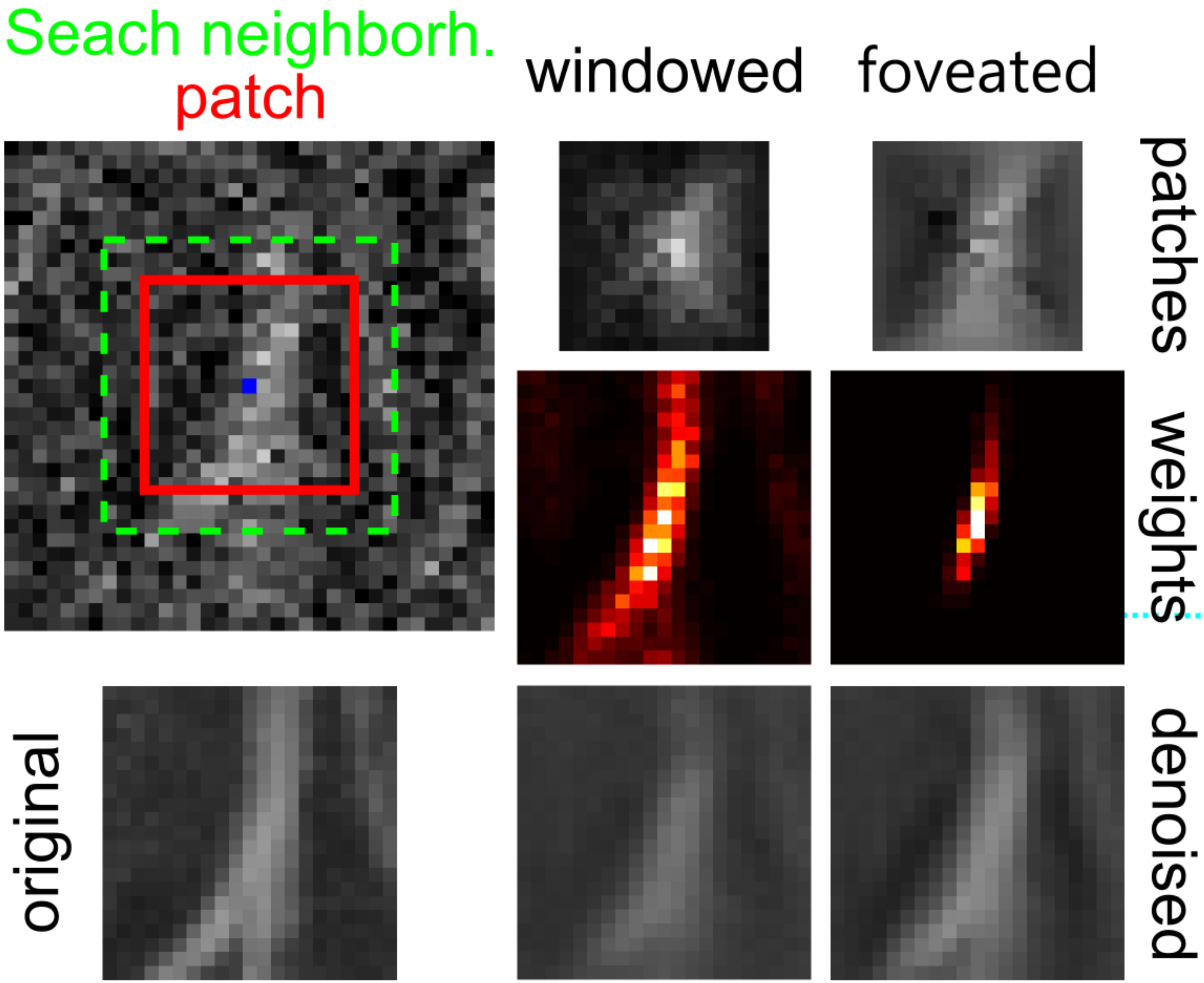
# Windowing vs Foveation



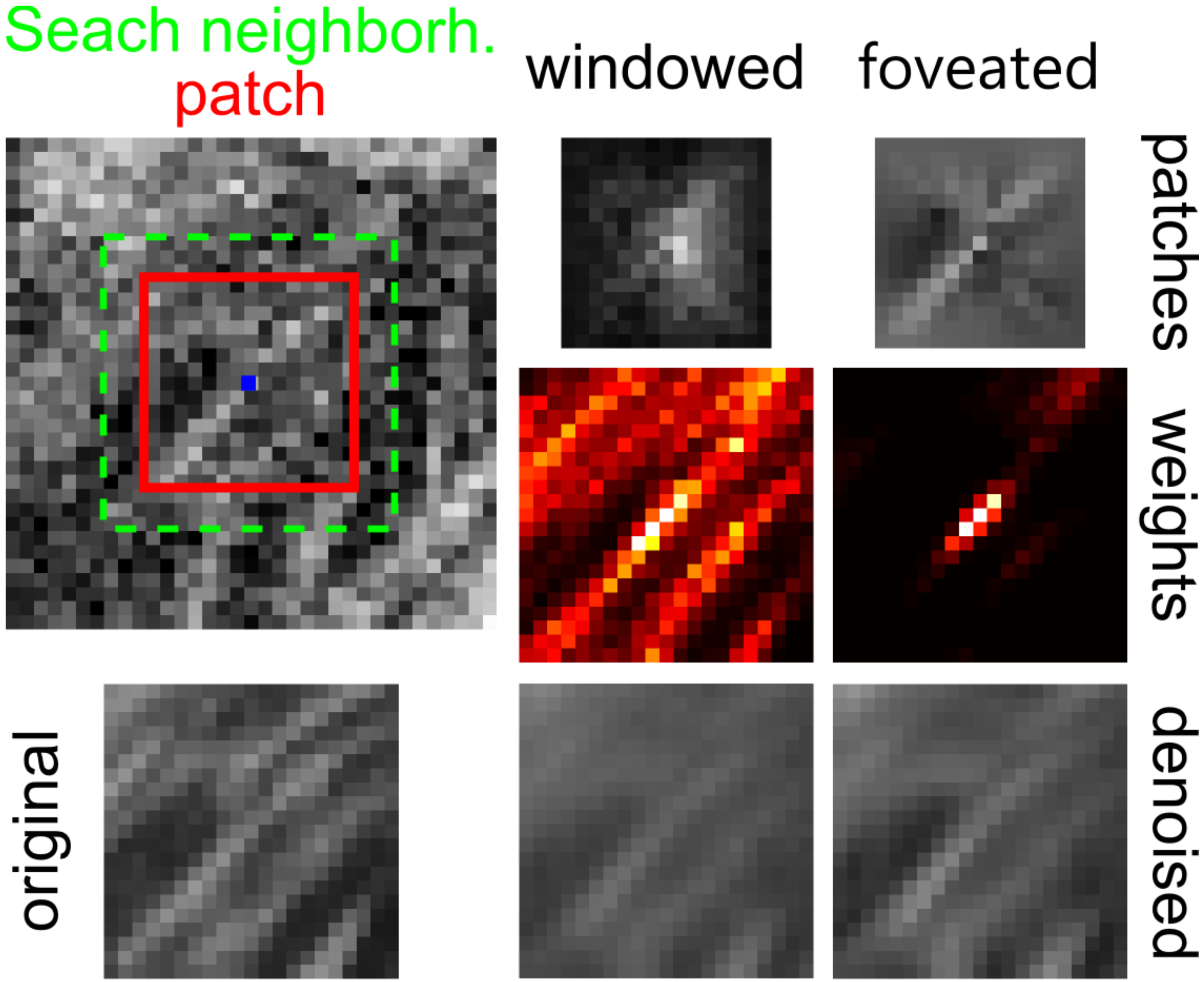
# Windowing vs Foveation



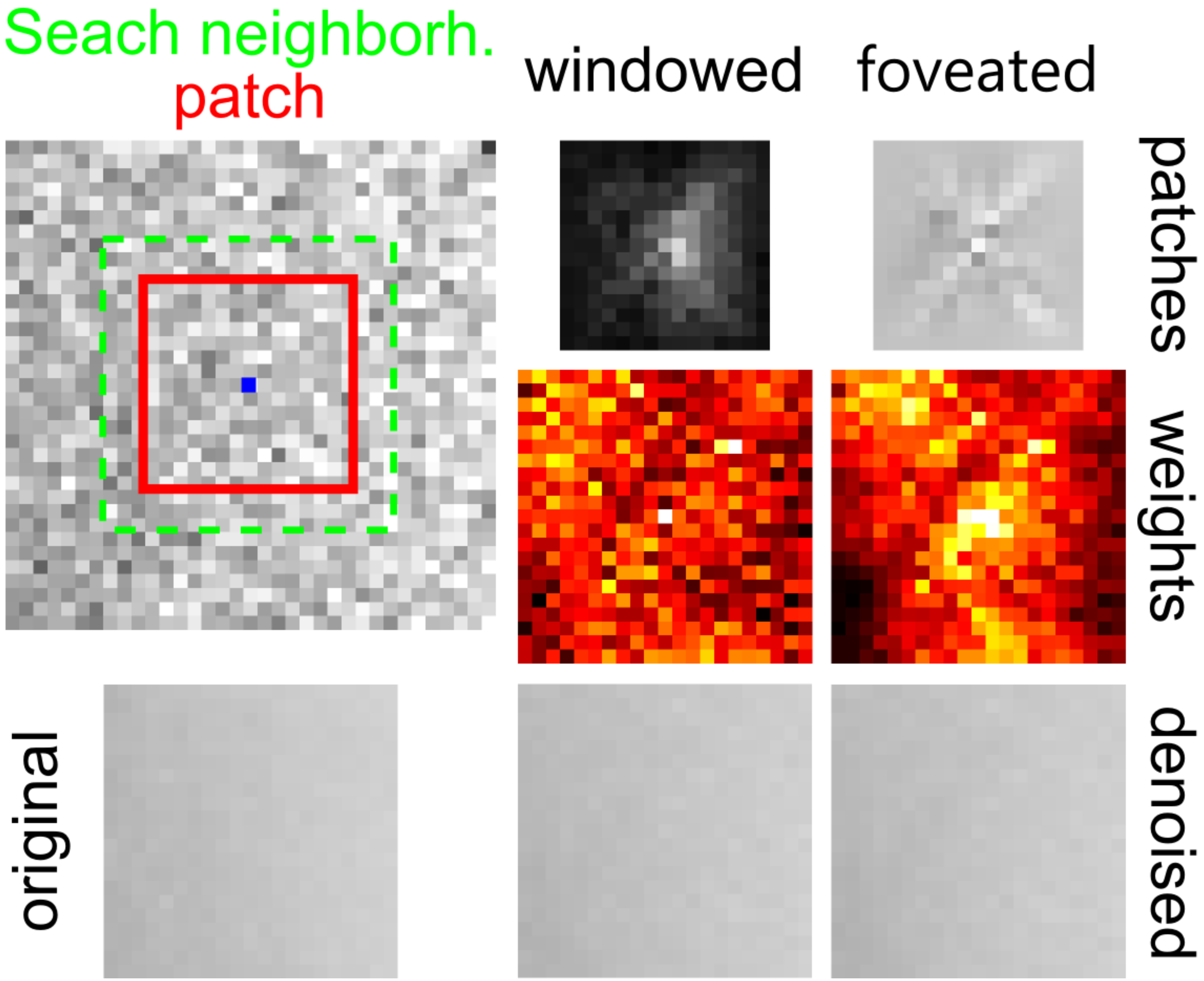
# Windowing vs Foveation



# Windowing vs Foveation



# Windowing vs Foveation

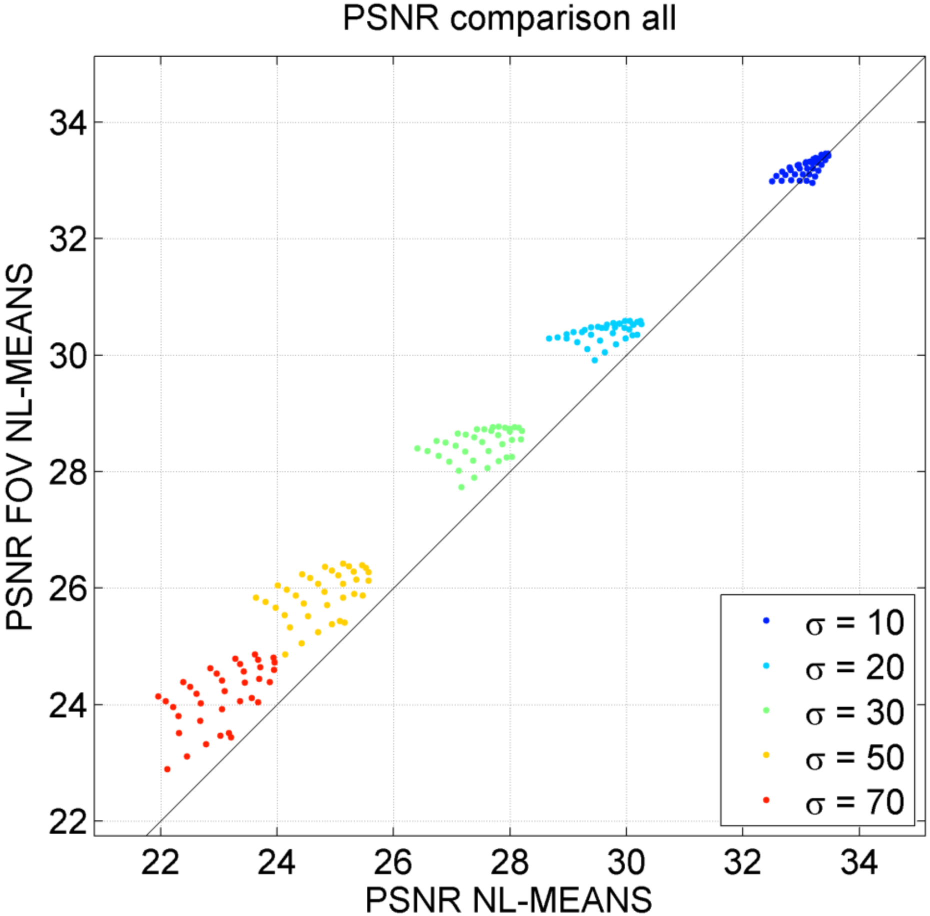


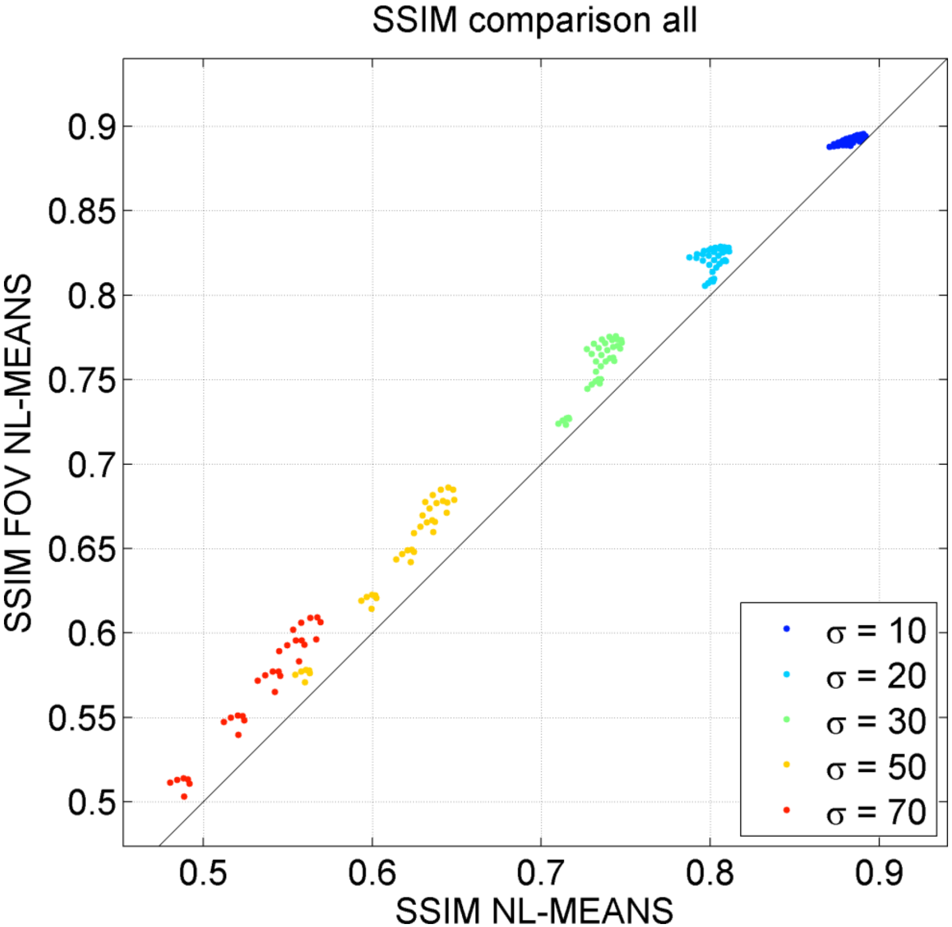
# Experiments (1/2)

Windowing vs Foveation

# NL-means vs. Foveated NL-means

- 7 test images $512 \times 512$ grayscale 8-bit images [0,255]
- noise standard deviation $\sigma = [10;\ 20;\ 30;\ 40;\ 50; 70]$
- $d$ and $d^{\mathrm{FOV}}$ are defined from the same kernel $\boldsymbol{k}$



# Noisy *Lena* $\sigma = 40$, $\mathbf{PSNR = 16.1\ dB}$

# NL-Means *Lena* $\sigma = 40$, $\mathbf{PSNR = 27.8\ dB}$

# Foveated NL-Means *Lena* $\sigma = 40, \mathbf{PSNR = 29.1\ dB}$

# NL-Means (green) vs Foveated NL-Means (red)

# Some Fragments of $150 \times 100$ pixels

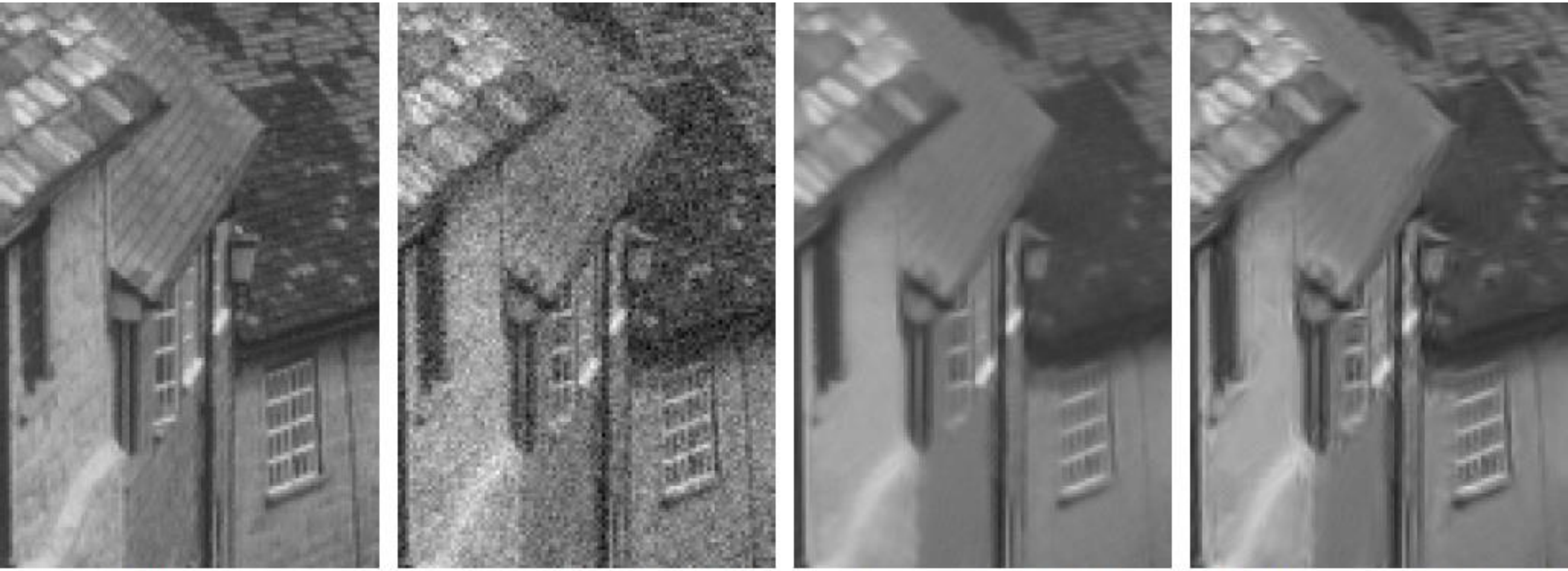
*Man*, original

noisy $\sigma = 10$ (28.14 0.680)

NL-means (32.24 0.867)

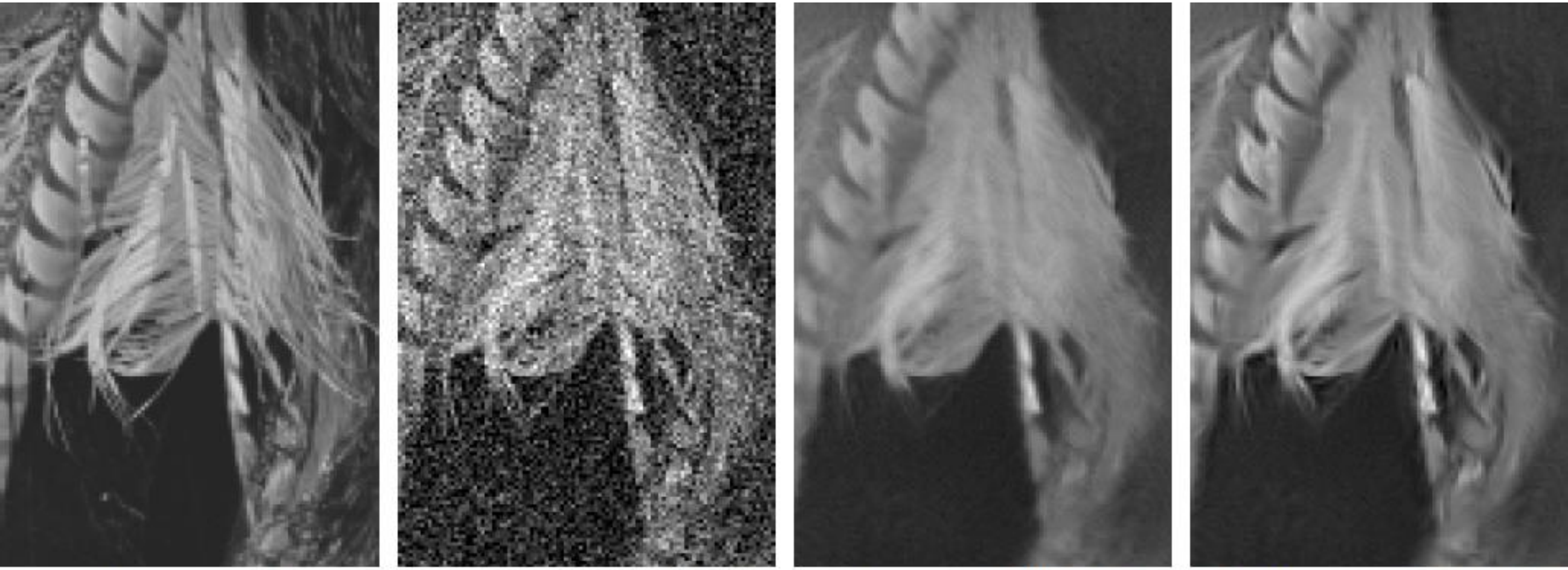
Fov. NL-means (32.84 0.890)

# Some Fragments of $150 \times 100$ pixels



*Hill*, original — noisy $\sigma = 15$ $(24.61\ 0.531)$ — NL-means $(30.22\ 0.779)$ — Fov. NL-means $(31.13\ 0.812)$

## Some Fragments of $150 \times 100$ pixels



*Man*, original | noisy $\sigma = 30$ (18.59 0.271) | NL-means (26.99 0.701) | Fov. NL-means (28.17 0.737)

# Some Fragments of $150 \times 100$ pixels

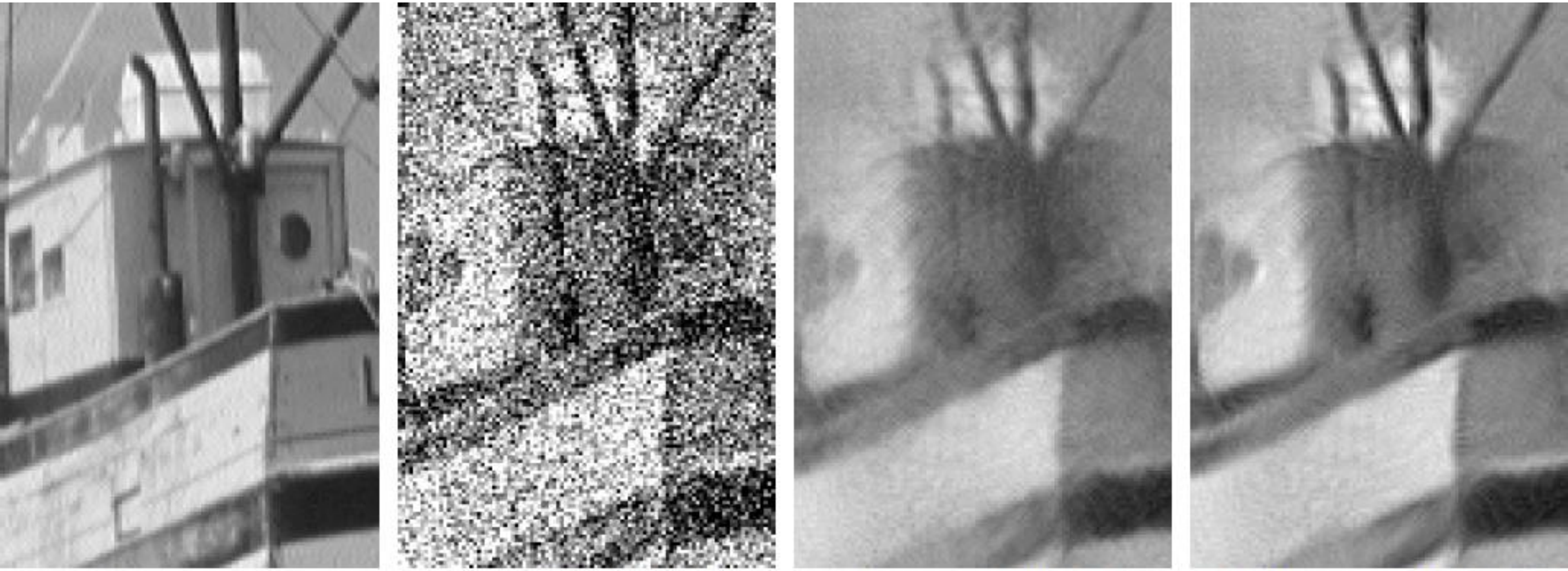
*Lena*, original

noisy $\sigma = 40$ (16.09 0.153)

NL-means (27.79 0.735)

Fov. NL-means (29.08 0.762)

# Some Fragments of $150 \times 100$ pixels



*Boats*, original — noisy $\sigma = 70$ (11.23 0.096) — NL-means (23.06 0.505) — Fov. NL-means (24.09 0.538)

# Discussion (1/2)

Windowing vs Foveation

# Discussion

- The **foveated self-similarity** a far more **effective regularization prior** (or descriptive model) for natural images than the conventional windowed self-similarity
  - Typical PSNR improvement in excess of 1 dB, especially at low SNR values.
  - Improvement in sharpness and visual perception (con.rmed by SSIM score).
- Foveated NL-Means is obtained as a direct modic..ation of the self-similarity measure within NL-means.
  - “Identical” computational complexity as standard NL-means.

## Windowing vs Foveation

- **Windowing** bears **no frequency selectivity** with respect to the patch content
  - typically the spatial autocorrelation of high-frequency subbands decays faster than that of low frequency components.
  - Likely variations in the high-frequencies may prevent the joint nonlocal filtering of otherwise mutually similar patches (+ variance);
  - The sensitivity with respect to variations in the low-frequencies is weakend by windowing (+ bias).
- In contrast, **foveation operators** provide a **compact multiscale representation** of each image patch:
- **Foveation can be interpreted as a conical sectioning of the scale-space** representation of an image

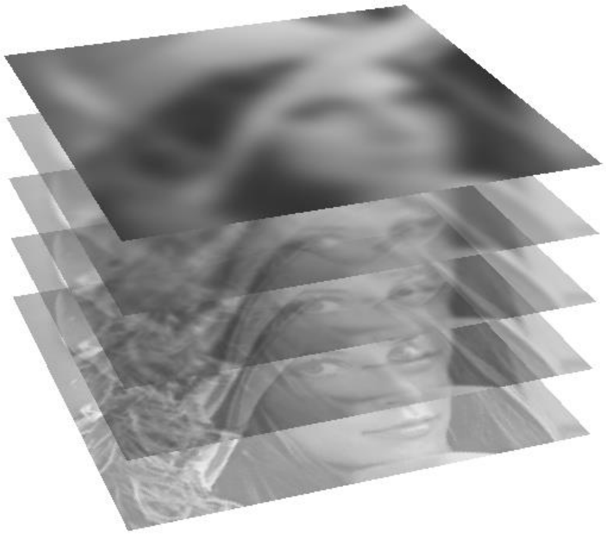
# Foveation as "conical" sectioning of the scale space

# Foveation as "conical" sectioning of the scale space

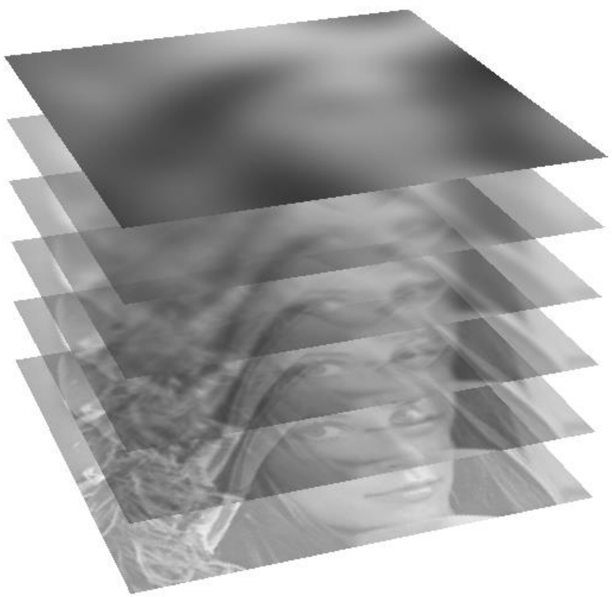
# Foveation as "conical" sectioning of the scale space

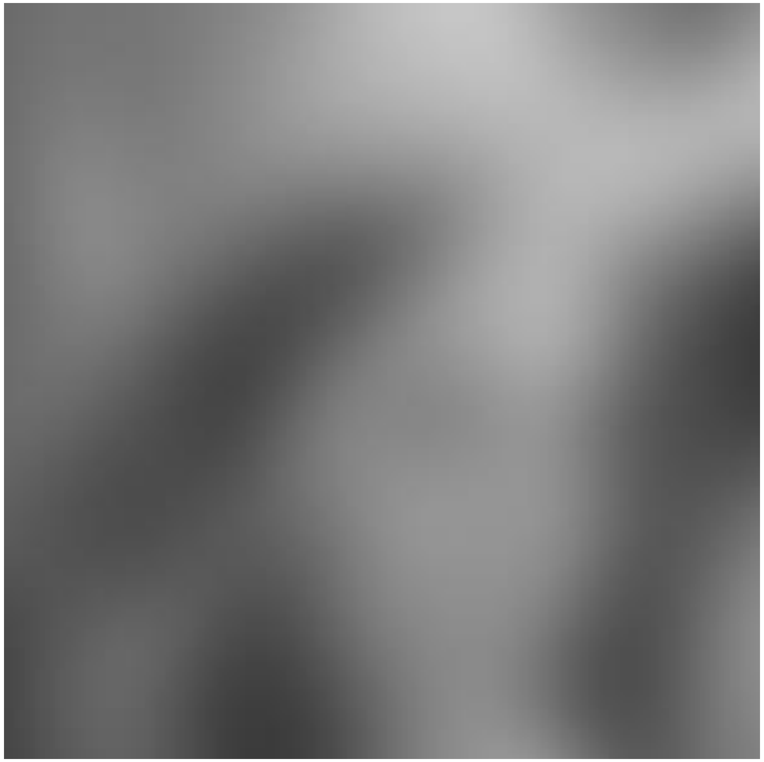
# Foveation as "conical" sectioning of the scale space

# Foveation as "conical" sectioning of the scale space

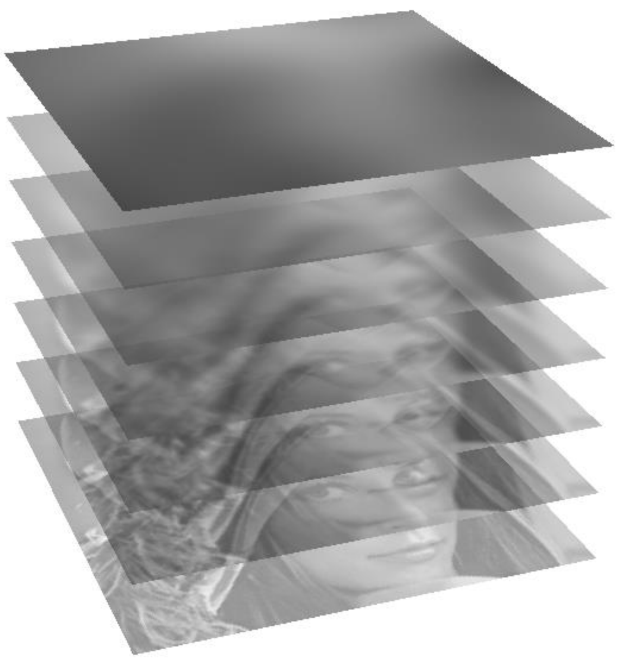


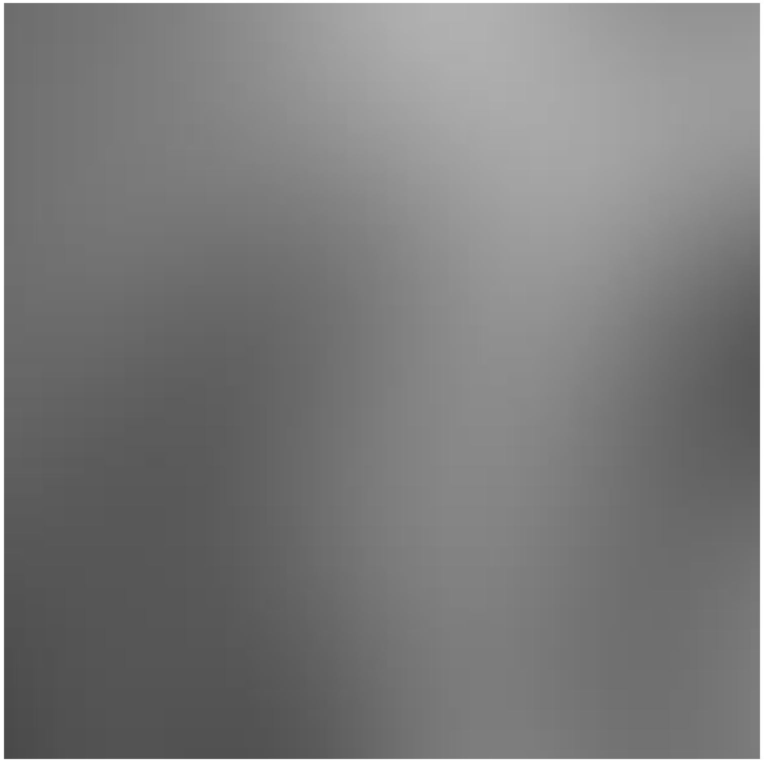
# Foveation as “conical” sectioning of the scale space



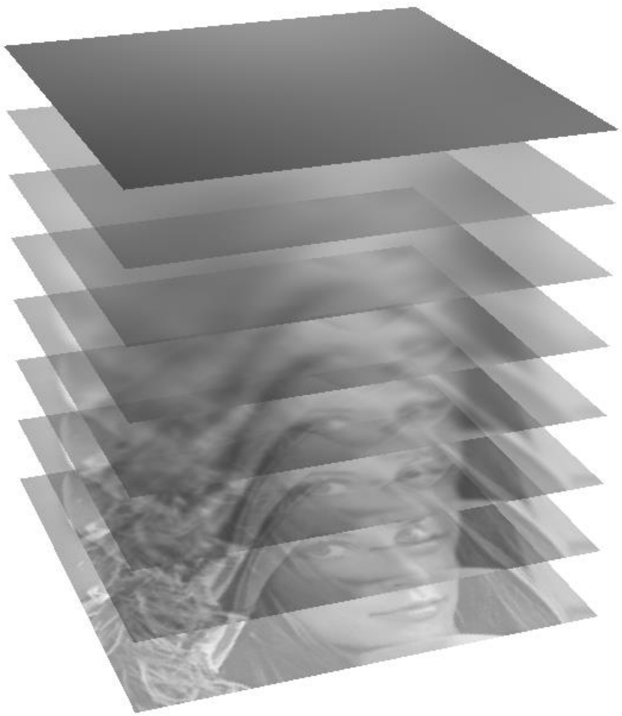


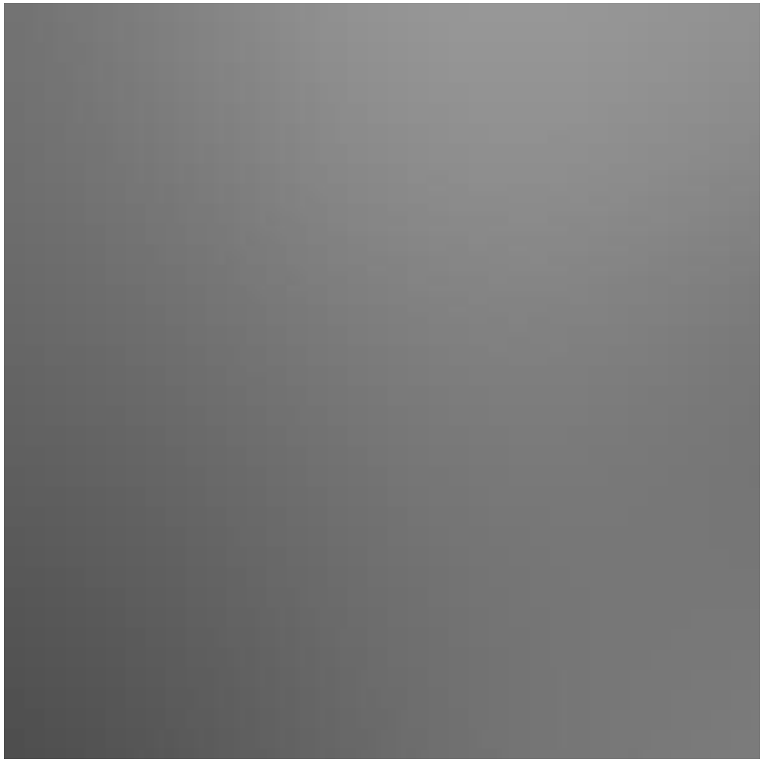
# Foveation as "conical" sectioning of the scale space



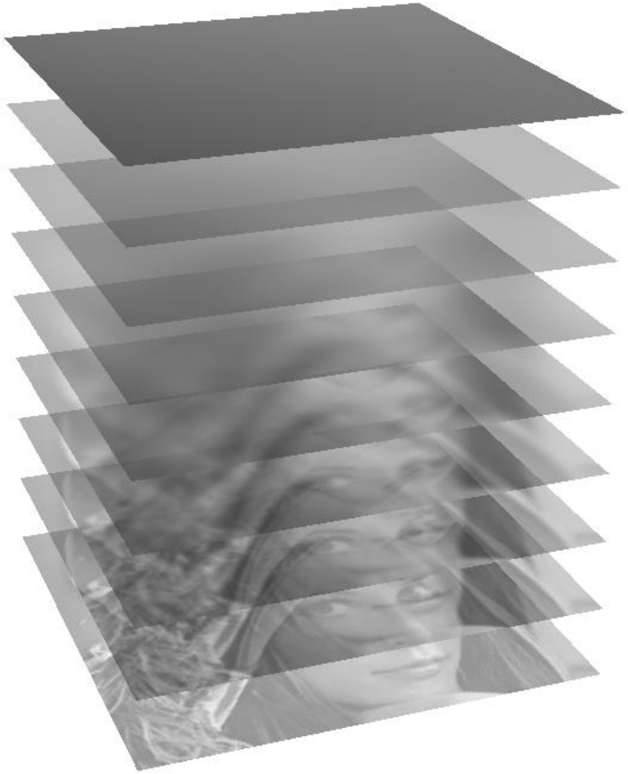


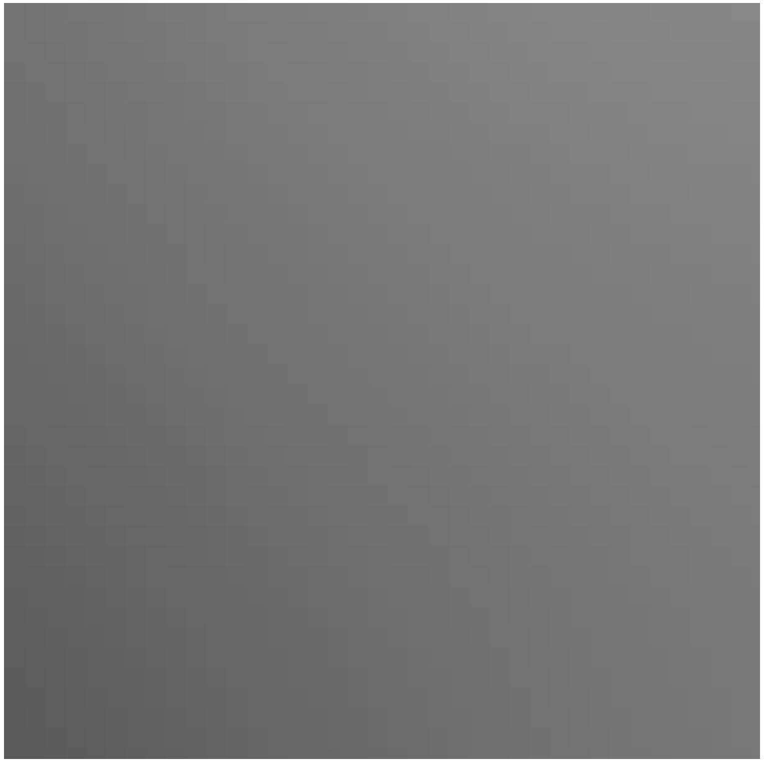
# Foveation as "conical" sectioning of the scale space



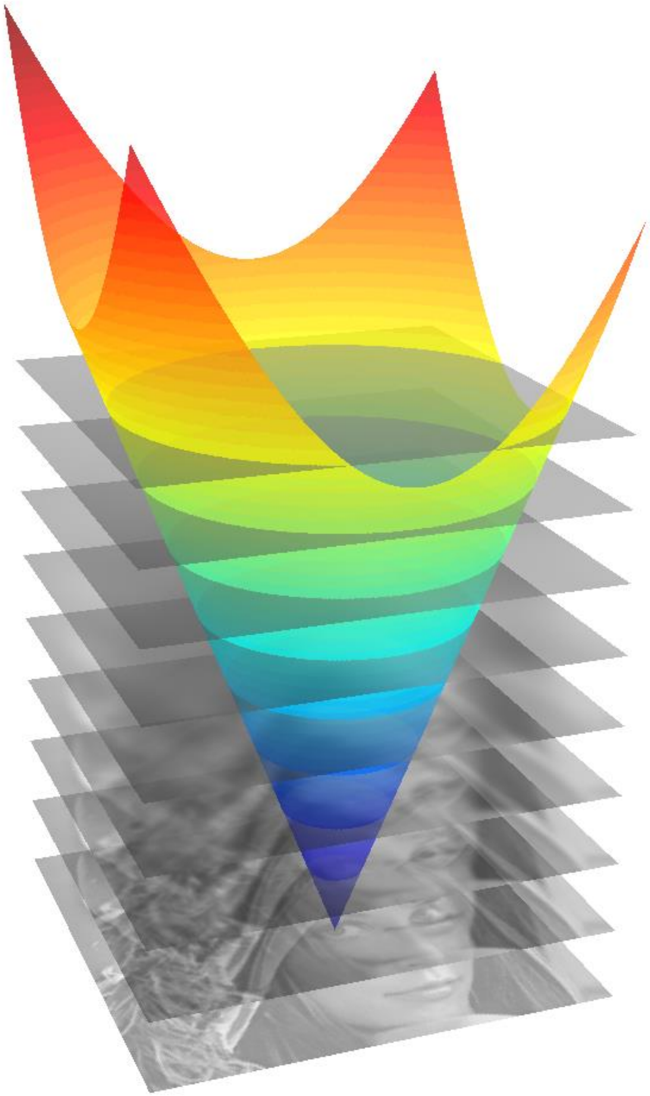


# Foveation as “conical” sectioning of the scale space

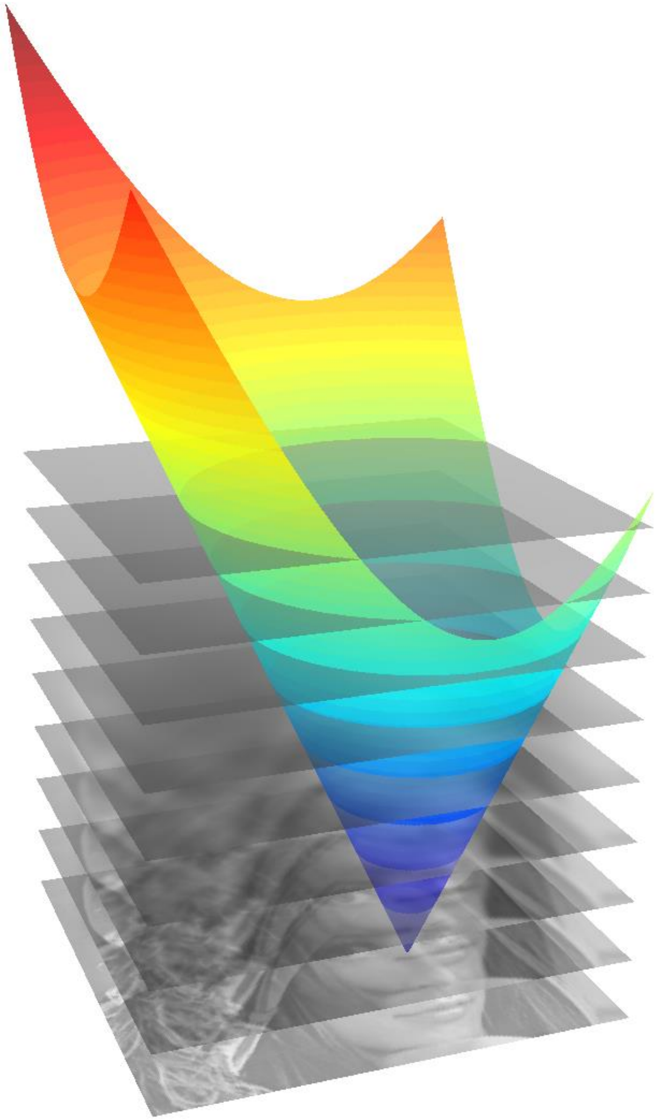




# Foveation as "conical" sectioning of the scale space



# Foveation as "conical" sectioning of the scale space



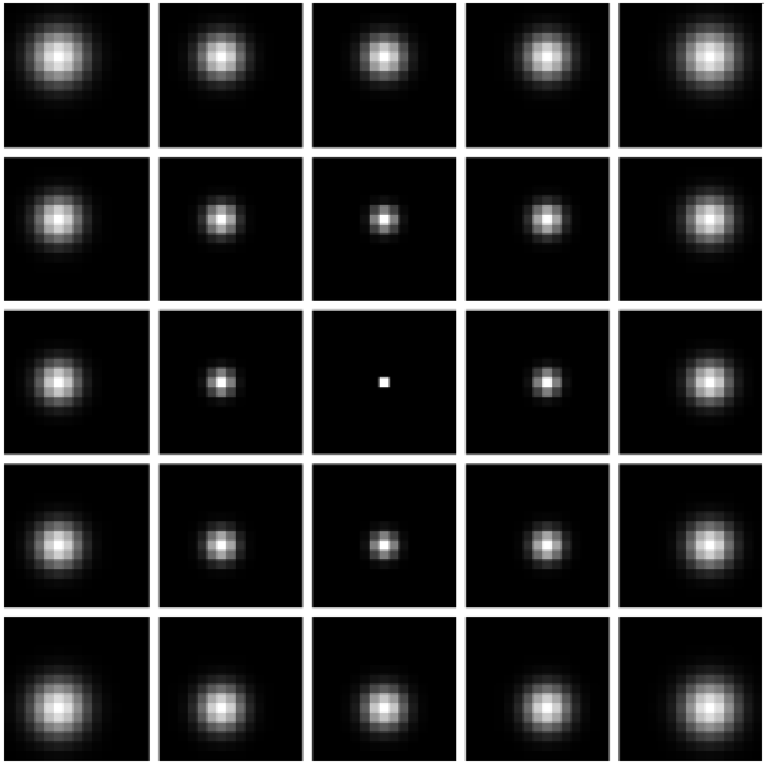
# Anisotropic Foveation

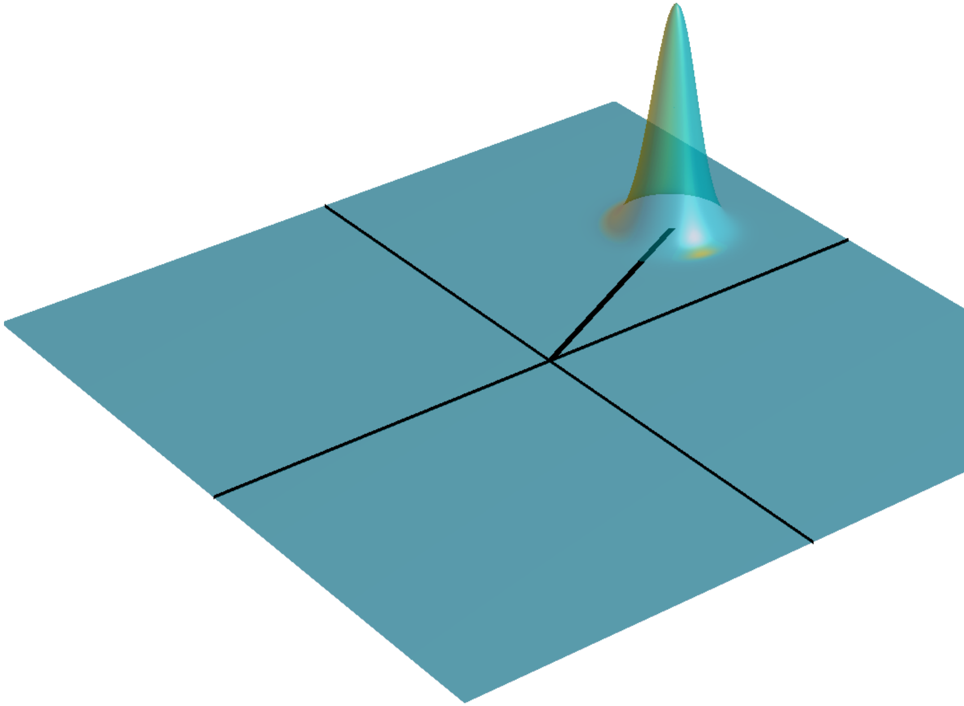
# Anisotropic Foveation Operators

- Generalization of the isotropic ones
- Use elliptical Gaussian PDF instead of the circularly-symmetric Gaussian PDF
- The covariance matrix of the Gaussians PDFs (yielding $\{v_u\}_{u \in U}$) depends on
  - The standard deviation of the kernel (given by $|u|$)
  - $\rho > 1$ that determines the elongation of the PDF,
  - $\theta$ an angular parameter (offset) that controls the orientation of the axes of the elliptical PDF

# Isotropic Foveation

- Use of circularly-symmetric Gaussian PSF

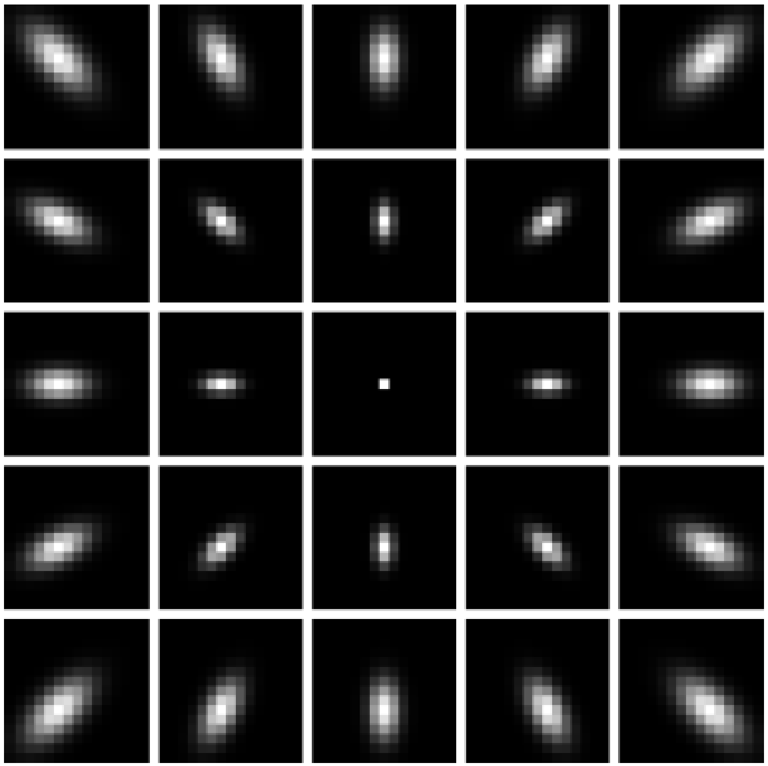


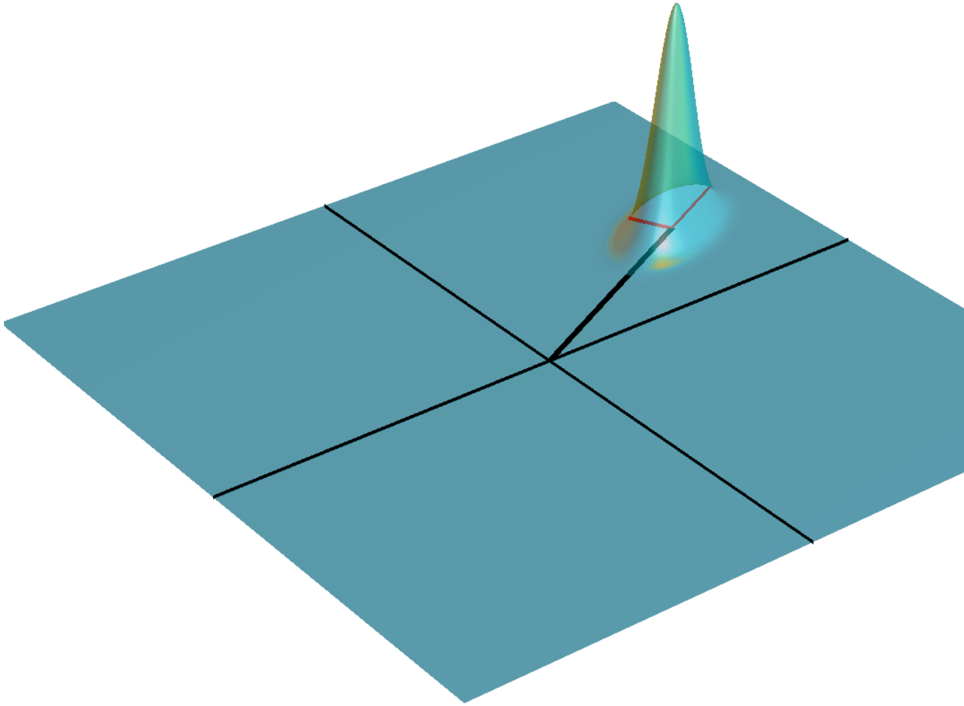


**Isotropic:** $\rho = 1, \ \theta =$ “any”

# Anistropic Foveation:

- Replace the circularly-symmetric Gaussian blur PSF with an elliptical Gaussian PSF

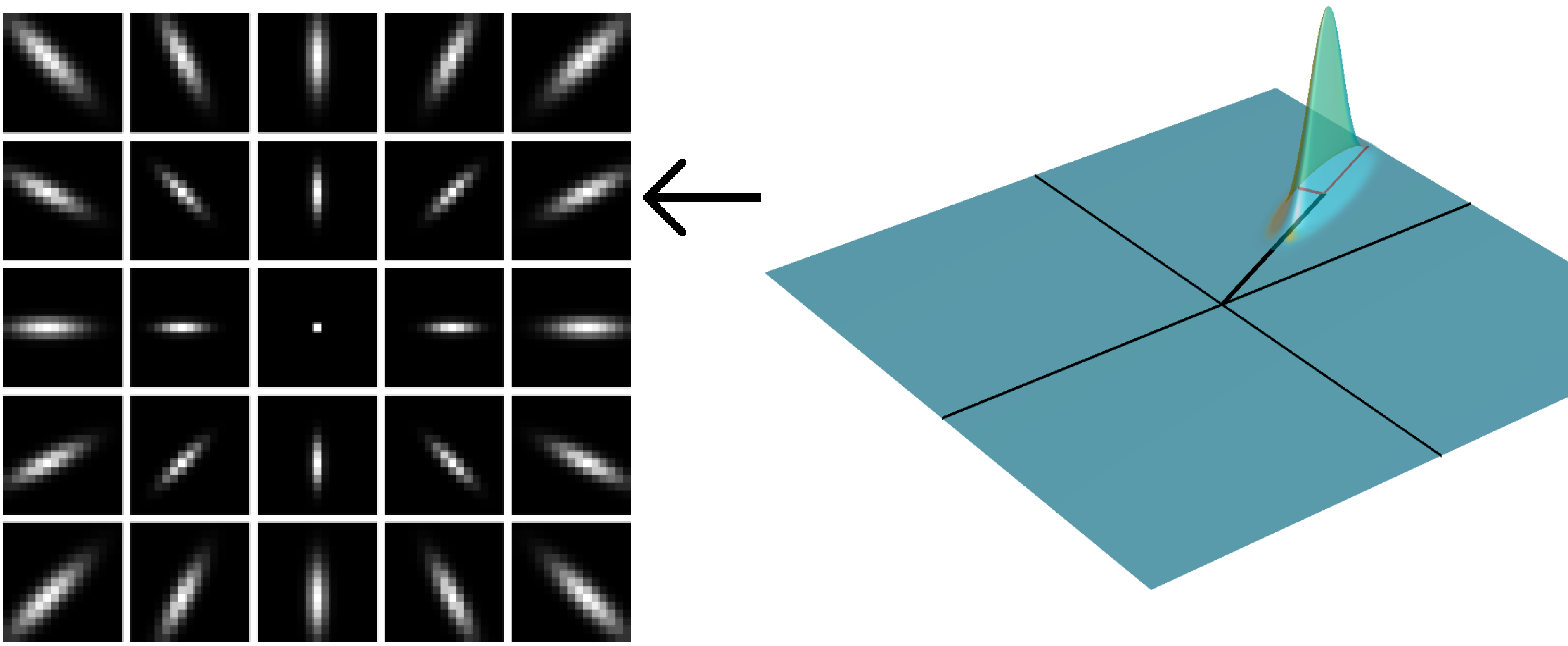




**Radial:** $\rho = 2$, $\theta = 0$

## Anistropic Foveation:

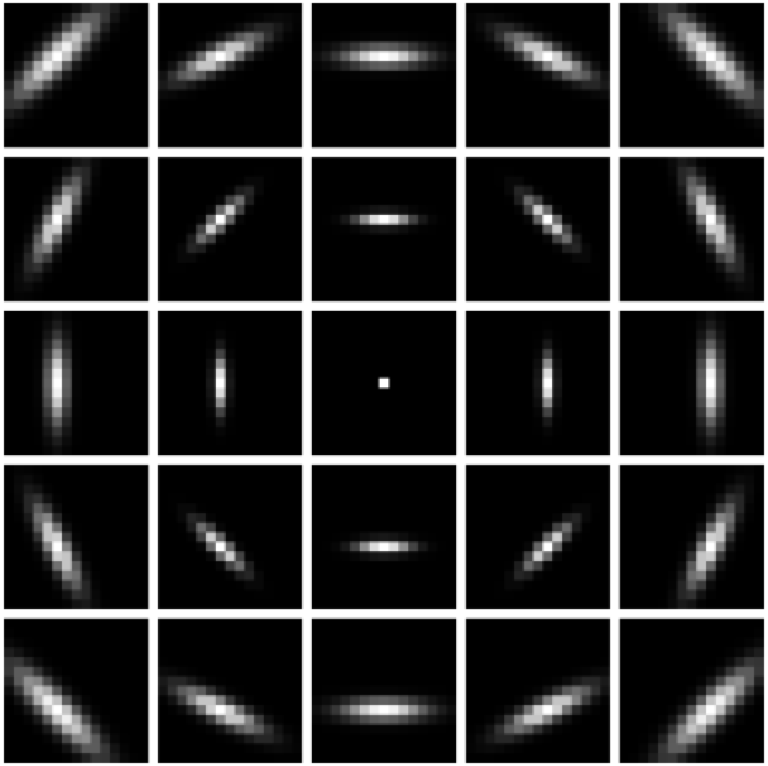
- Replace the circularly-symmetric Gaussian blur PSF with an elliptical Gaussian PSF

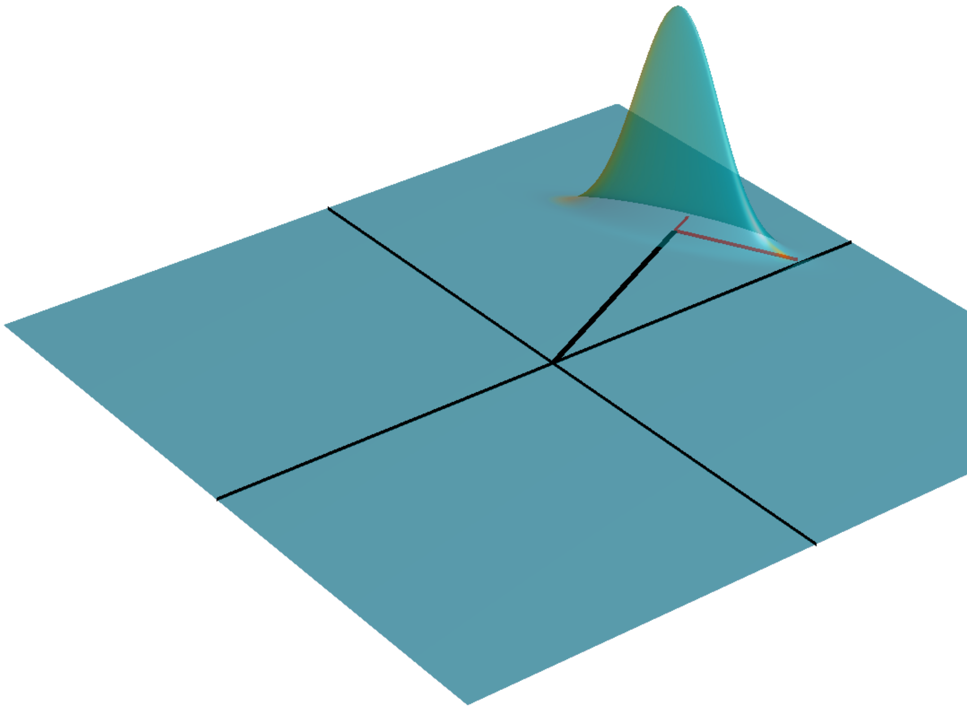


**Radial:** $\rho = 4$, $\theta = 0$

# Isotropic Foveation

- Use of circularly-symmetric Gaussian PSF



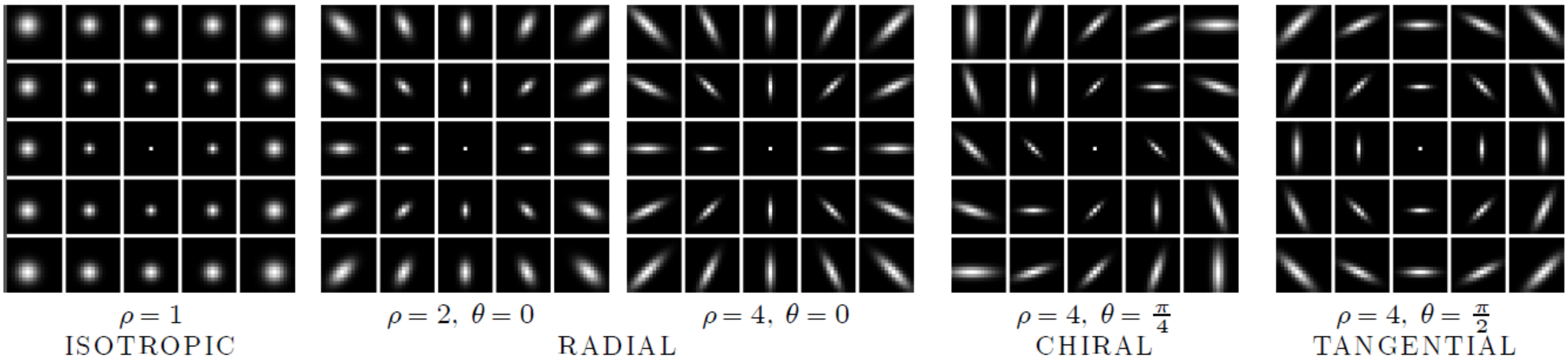


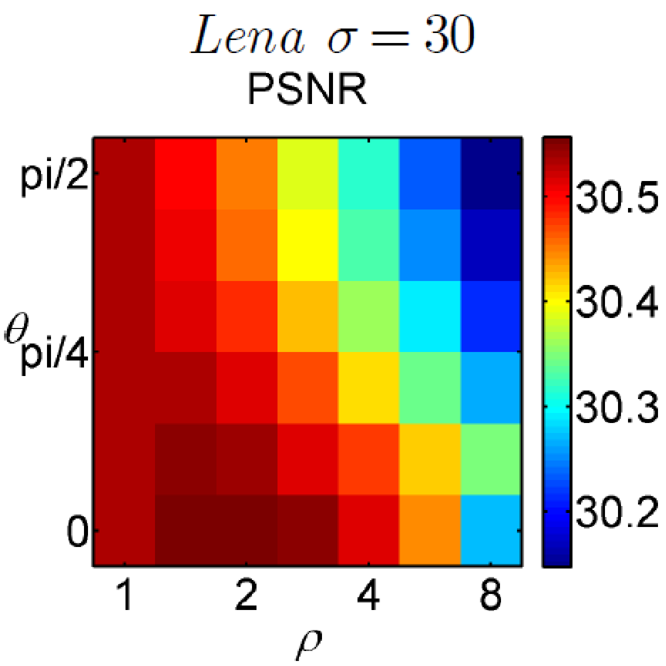
**Tangential:** $\rho = 4,\ \theta = \frac{\pi}{2}$

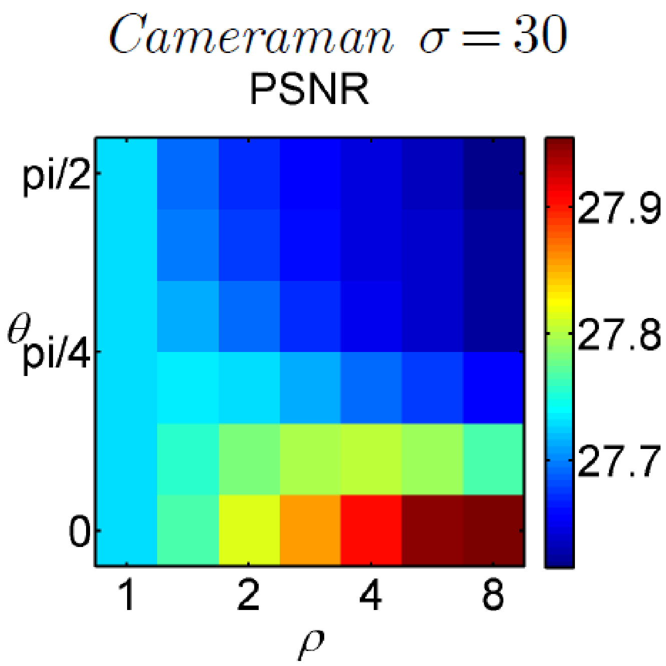
# Experiments (2/2)

Anisotropic Foveation

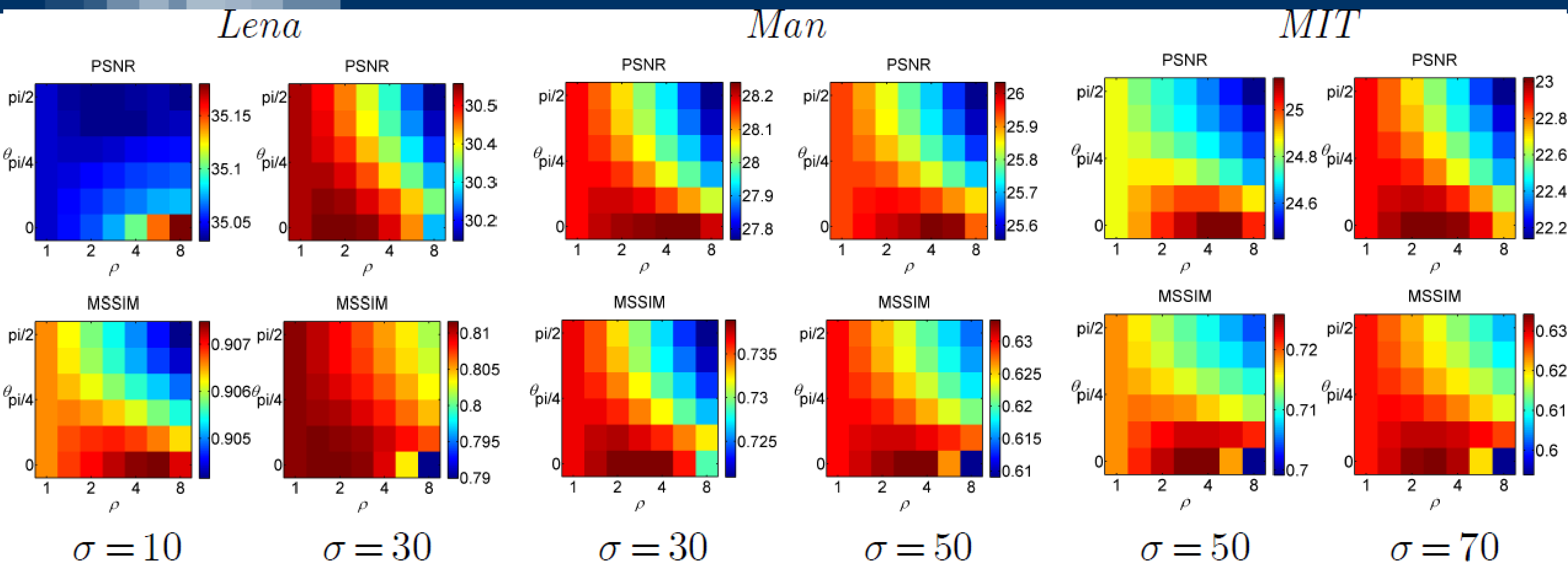
# PSNR for Anisotrpic Foveation

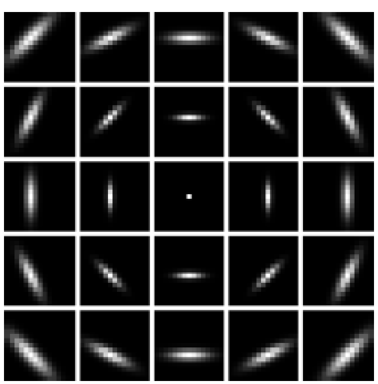


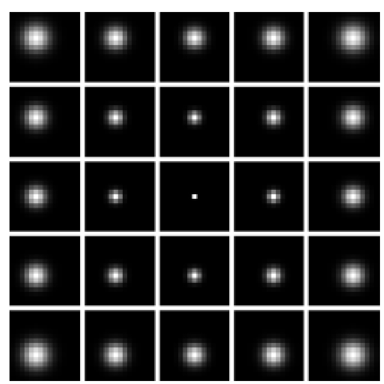



# PSNR and SSIM for Anisotrpic Foveation

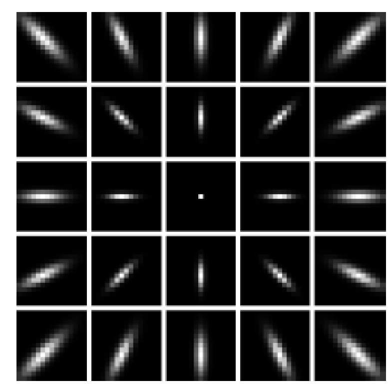


$\rho = 4, \ \theta = \frac{\pi}{2}$
TANGENTIAL (**BAD**)



$\rho = 1$
ISOTROPIC (**OK**)



$\rho = 4, \ \theta = 0$
RADIAL (**BETTER**)

# Radial (RED), Isotropic (YELLOW), Tangential (GREEN)

# Discussion (2/2)

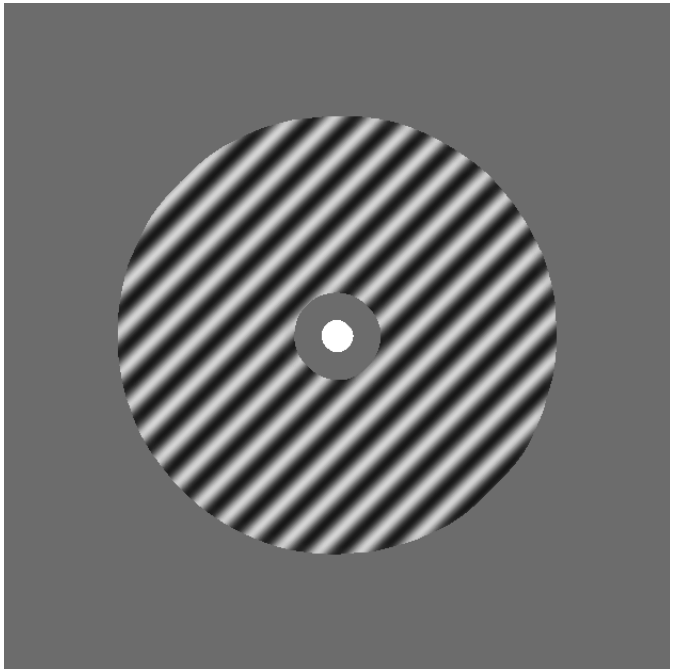
Anisotropic Foveation

## Comments

- **Radial foveation** is a more effective regularity assumption the isotropic foveation or the windowing
  - It preserves the substantial edge structure, since pixels are blurred along the edge rather than across the edge. Similar arguments lead to Anisotropic NL-means (Maleki et al 2013)
- The performance gap between radial and isotropic foveation is less substantial than foveation against windowing
- The improvements achived by radial foveation operators recall the *radial orientation bias of human visual system* (Sasaki et al, 2006, Freeman et al, 2011)

# Orientation preference in the human visual system

- The oriented stimuli wich is displayed while rotating
- The brain activity in an annular region of V1 is monitored
- There is a stronger response when the stimuli orientation agree with the angular position of the retinally mapped voxels of V1 (Freeman et al, 2011)
- Patches blurred by radial foveation operators preserves edges and sharp details directed towards the patch center
- PSF layout recalls the orientation preference

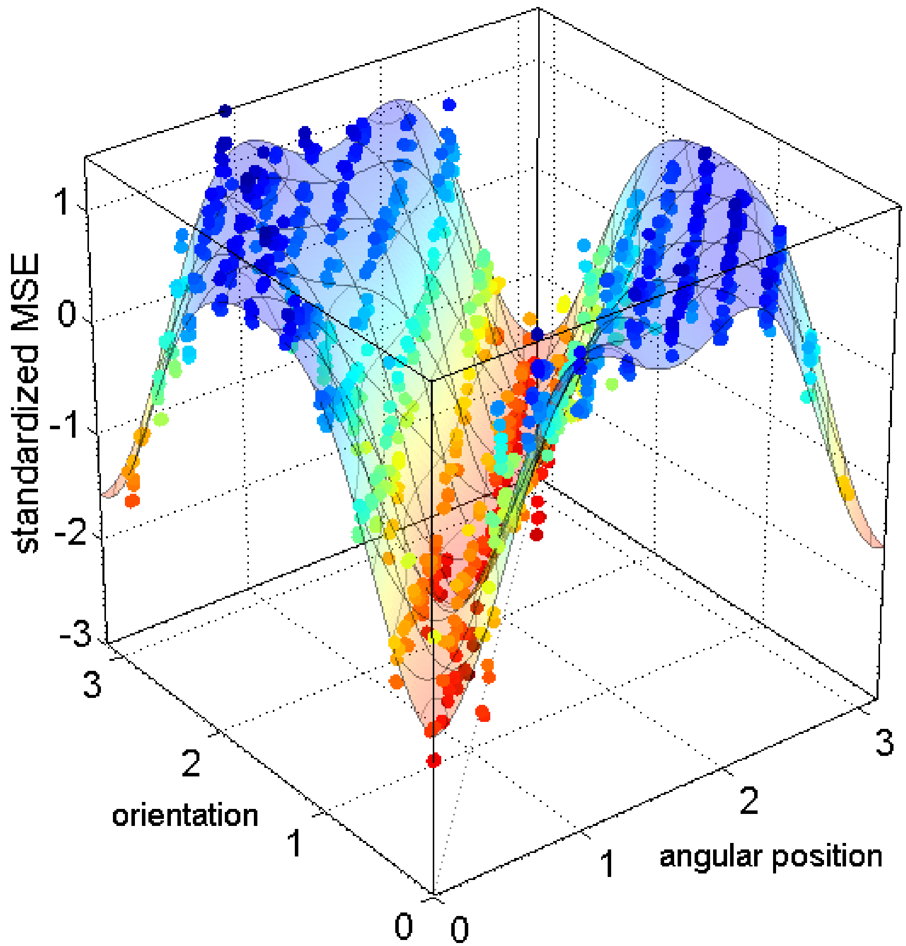


# Orientation preference in human primary visual cortex

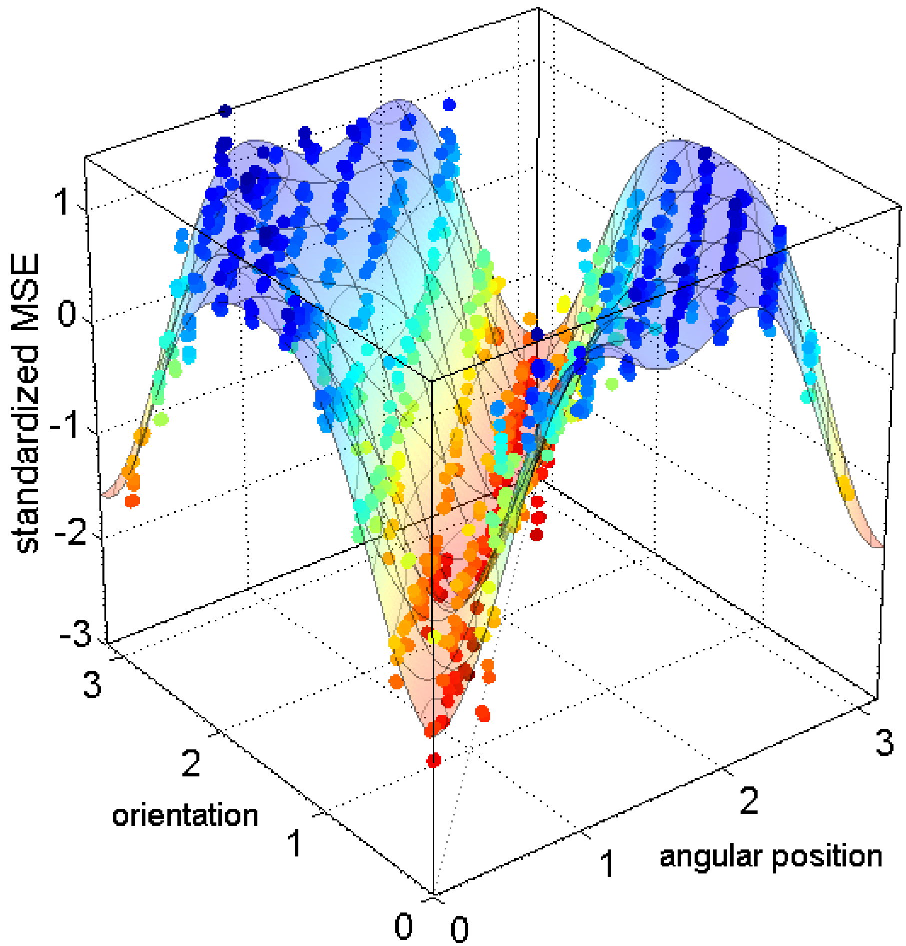
- See Figure 3 in

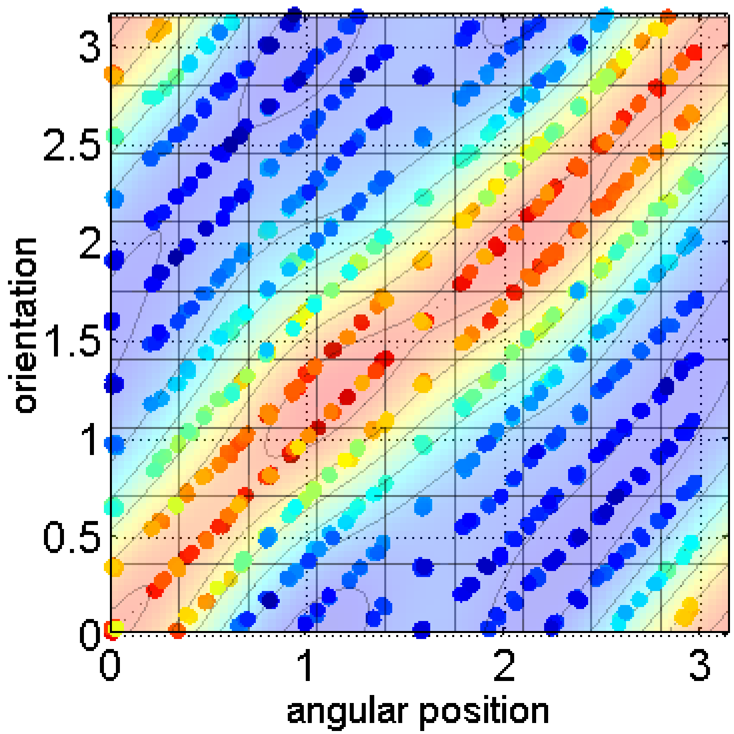
**J. Freeman, G. J. Brouwer, D. J. Heeger, E. P. Merriam, "Orientation Decoding Depends on Maps, Not Columns," J. Neurosc., March 2011.**

# Orientation Preference from NL-means



# Orientation Preference from NL-means



# Conclusions and Ongoing Works

- Radial foveation is a more effective regularity assumption the isotropic foveation or the windowing
- Anisotropic foveation exploits an essential principle from human vision. Agrees with recent findings in neuroscience.

- **Ongoing Works**
  - Extension to multiscale nonlocal transform-domain filtering.
  - Analysis of foveated self-similarity in the context of natural image statistics.
  - Investigate connection with transaccadic integration

# References

- Source codes and references online at
  http://www.cs.tut.fi/~foi/FoveatedNL/
  http://home.deib.polimi.it/boracchi/Projects/Foveation.html

- **Papers**
  - **Anisotropically Foveated Nonlocal Image Denoising**
    Alessandro Foi, Giacomo Boracchi *in Proceedings of ICIP 2013, IEEE International Conference on Image Processing, September 15-18 2013, Melbourne, Australia*
  - **Foveated self-similarity in nonlocal image filtering**
    Alessandro Foi, Giacomo Boracchi *Proc. IS&T / SPIE Electronic Imaging 2012, Human Vision and Electronic Imaging XVII (2012), 12 pages, January 2012 Burlingame (CA), USA. doi:10.1117/12.912217*
  - *Journal paper is on the way*