NONLOCAL FOVEATED PRINCIPAL COMPONENTS

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

Patch foveation corresponds to a spatially variant representation where the center of the patch is sharp while the periphery is blurred. This mimics the non-uniformity of the human visual system, whose acuity is maximal at the fixation point (imaged by the fovea, i.e. the central part of the retina) and low at the periphery of the visual field.

We introduce patch foveation for patch clustering in dictionary learning. In particular, we consider principal components learned from clusters of foveated patches extracted from natural images corrupted by additive noise. Experiments demonstrate that the first few foveated principal components provide a better representation of the actual (non-foveated) image than the usual principal components learned from clusters of patches or windowed patches. These new results confirm the effectiveness of patch foveation as regularization and preconditioning prior when processing natural images.

Index Terms— Dictionary learning, Principal Components, Nonlocal Similarity, Denoising, Patches, Foveation.

1. INTRODUCTION

During the last decade there has been a growing interest towards dictionaries or transforms learned from the data. In particular, for the purpose of image filtering, we witnessed a systematic trend towards online learning from structured subsets often extracted from the data to be processed. Just like adaptively learned global dictionaries typically provide more efficient (i.e. sparse) representations than non-adaptive dictionaries or fixed transforms [1], semi-local dictionaries or transforms learned from multiple clusters of image patches typically outperform the global ones [2]. This observation explains the success of current state-of-the-art image denoising filters such a BM3D-SAPCA [3] or PLOW [4].

A crucial aspect in the above works concerns the selection of the cluster used for learning the dictionary. A patch-similarity metric is usually utilized in order to define the cluster around a given "reference" or "centroid" patch. The methods rely on the assumption that natural images contain a large number of mutually similar patches at different locations, hence meaningful clusters are likely to be found. Patch

similarity is typically assessed through the Euclidean or a weighted-Euclidean distance of the pixel intensities.

In [5], we first introduced the foveated patch distance as an alternative to the weighted-Euclidean distance in nonlocal imaging, leading to the concept of foveated self-similarity. Patch foveation is performed through a suitably designed 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. This approach is inspired by the human visual system (HVS), which features spatially variant properties [6]: 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 [7, 8] we generalized the approach by introducing anisotropic foveation operators, realized by means of directional PSFs.

When plugged into a classical pointwise nonlocal filter such as the NL-means [9], the foveated distance – especially the one induced by radially anisotropic foveation operators – provides essential improvement in denoising quality with respect to the conventional windowed distance. Such improvement is verified in terms of both objective criteria and visual appearance, particularly due to better contrast and sharpness.

In this work we explore the impact of foveation in the identification of patch clusters for the purpose of adaptive dictionary learning. For the sake of simplicity and clarity, we restrict ourselves to the principal components (PC), as eigenvectors of the second-moment matrix of a patch cluster. Patch foveation operates jointly on two aspects: first, it replaces the patches with their foveated counterparts; second, the foveated distance (i.e. the Euclidean distance between the foveated patches) defines the patch similarity, which induces the weights in the second-moment matrix [10]. We call this special form of weighted PCs the *nonlocal foveated principal components*, or, simply, the *foveated principal components*. Since the components are learned on foveated patches, in the analysis they have to be composed with the foveation operator; likewise, in the synthesis, they are composed with the inverse foveation operator. The reported experimental results demonstrate that the foveated PCs provide much better reconstruction of the overall image data than standard PCs based on a conventional windowed patch distance, in spite of the

---

fact that the analysis is carried out on partly blurred patches. These results are consistent with our previous findings on the Foveated NL-means and confirm the useful interplay between foveation and nonlocal image processing. On a very broad sense, this newly introduced use of foveation can be interpreted as a form of preconditioning for the cluster selection, and consequently for the dictionary learning.

2. PRELIMINARIES

2.1. Observation Model

We consider noisy grayscale images $z : X \to \mathbb{R}$ 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), $y : X \to \mathbb{R}$ is the unknown original noise-free image, and $\eta : X \to \mathbb{R}$ is i.i.d. Gaussian white noise, $\eta(\cdot) \sim \mathcal{N}(0, \sigma^2)$.

2.2. Patches

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

Given a non-negative *windowing kernel* $\mathbf{k}$ defined over U, we can naturally define the *windowed patches* $\mathbf{z}_x^{\text{WIN}} = \mathbf{z}_x \sqrt{\mathbf{k}}$.

Foveated patches are denoted as $\mathbf{z}_x^{\text{FOV}}$, and are obtained through the following foveation operators.

2.3. Foveation Operator

A foveated patch $\mathbf{z}_x^{\text{FOV}}$ and an anisotropic foveation operator $\mathcal{F}_{\rho,\theta}$ take the generic form

$$\mathbf{z}_x^{\text{FOV}}(u) = \mathcal{F}_{\rho,\theta}[z,x](u) = \sum_{\xi \in \mathbb{Z}^2} z(\xi + x)\, v_u^{\rho,\theta}(\xi - u), \quad (2)$$

where $u \in U$ and $v_u^{\rho,\theta}$ is a bivariate elliptical Gaussian PSFs producing the blur. The blur increases with $|u|$. The parameter $\rho > 0$ determines the elongation of the PSF (as the ratio between its primary and secondary axis), and the angular-offset parameter θ controls the relative orientation of the PSFs with respect to the center of the patch. Specifically, $\theta = \vartheta - \angle u$, where ϑ is the absolute orientation of the primary axis and $\angle u$ is the angle of the segment between u and the center of the patch. *Radial* foveation operators are obtained for $\theta = 0$ and $\rho > 1$, i.e. when the major axes of the PSFs are directed along the meridian lines toward the patch center. Due to limitation of space, we refer the reader to [5],[8] for illustrations of foveated patches and foveation operators.

2.4. Patch Distances

We consider three difference distances between patches centered at generic positions x_1 and x_2: the usual Euclidean distance,

$$d(x_1, x_2) = \frac{1}{|U|} \|\mathbf{z}_{x_1} - \mathbf{z}_{x_2}\|_2^2, \quad (3)$$

the windowed quadratic distance,

$$d^{\text{WIN}}(x_1, x_2) = \frac{1}{\|\mathbf{k}\|_1} \left\| \mathbf{z}_{x_1}\sqrt{\mathbf{k}} - \mathbf{z}_{x_2}\sqrt{\mathbf{k}} \right\|_2^2, \quad (4)$$

and the foveated distance

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

If we consider the patch center as a fixation point, the foveated distance mimics the inability of the HVS to perceive details at the periphery of the center of attention. A constrained design of the foveation operators [5, 8] guarantees that whenever the difference between two patches can be modelled as white noise, as it is the case when the two patches do not exhibit structured differences, foveated distance and windowed distance are equivalent. 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 sharper localized responses.

3. NONLOCAL FOVEATED PRINCIPAL COMPONENTS

For any given patch $\mathbf{z}_{\bar{x}}$ (centroid), we construct a distance-weighted cluster by considering all patches $\{\mathbf{z}_j\}_{j=1}^N$ found in a neighborhood of $\bar{x}$ (potentially the whole image) and their foveated distance from $\mathbf{z}_{\bar{x}}$, i.e. $d^{\text{FOV}}(\bar{x}, x_j)$, $j = 1, \ldots, N$. We then construct the sample weighted second-moment matrix

$$C_{\bar{x}} = \left[\overrightarrow{\mathbf{z}_1^{\text{FOV}}} \cdots \overrightarrow{\mathbf{z}_N^{\text{FOV}}} \right] W_{\bar{x}} \left[\overrightarrow{\mathbf{z}_1^{\text{FOV}}} \cdots \overrightarrow{\mathbf{z}_N^{\text{FOV}}} \right]^T, \quad (6)$$

where $\overrightarrow{\mathbf{z}_j^{\text{FOV}}}$ denotes the foveated patch $\mathbf{z}_j^{\text{FOV}}$ represented as column vector, and where the diagonal matrix $W_{\bar{x}}$ with exponential weights, $W_{\bar{x}} = \text{diag}(w(\bar{x}, x_1), \ldots, w(\bar{x}, x_N))$, adjusts the contribution of each patch to the overall matrix depending on the foveated distance $d^{\text{FOV}}(\bar{x}, x_j)$, $j = 1, \ldots, N$:

$$w(\bar{x}, x_j) = e^{-\frac{d^{\text{FOV}}(\bar{x}, x_j)}{\sigma^2}} \Big/ \sum_{i=1}^N e^{-\frac{d^{\text{FOV}}(\bar{x}, x_i)}{\sigma^2}}. \quad (7)$$

When a weight $w(\bar{x}, x_j)$ is zero, it is equivalent to excluding the corresponding patch $\mathbf{z}_j^{\text{FOV}}$ from the cluster. Due to the exponential decay of the weights (7), this situation is practically met whenever $d^{\text{FOV}}(\bar{x}, x_j)$ is large.

Eigenvalue decomposition of $C_{\bar{x}}$ yields

$$U_{\bar{x}}^T C_{\bar{x}} U_{\bar{x}} = S_{\bar{x}}$$

where $U_{\bar{x}}$ is an orthonormal matrix composed of eigenvectors and $S_{\bar{x}}$ is a diagonal matrix of eigenvalues in decreasing order of magnitude. Each column of $U_{\bar{x}}$ is an eigenvector, i.e. a foveated PC.

PCs	Windowed PCs	Foveated PCs
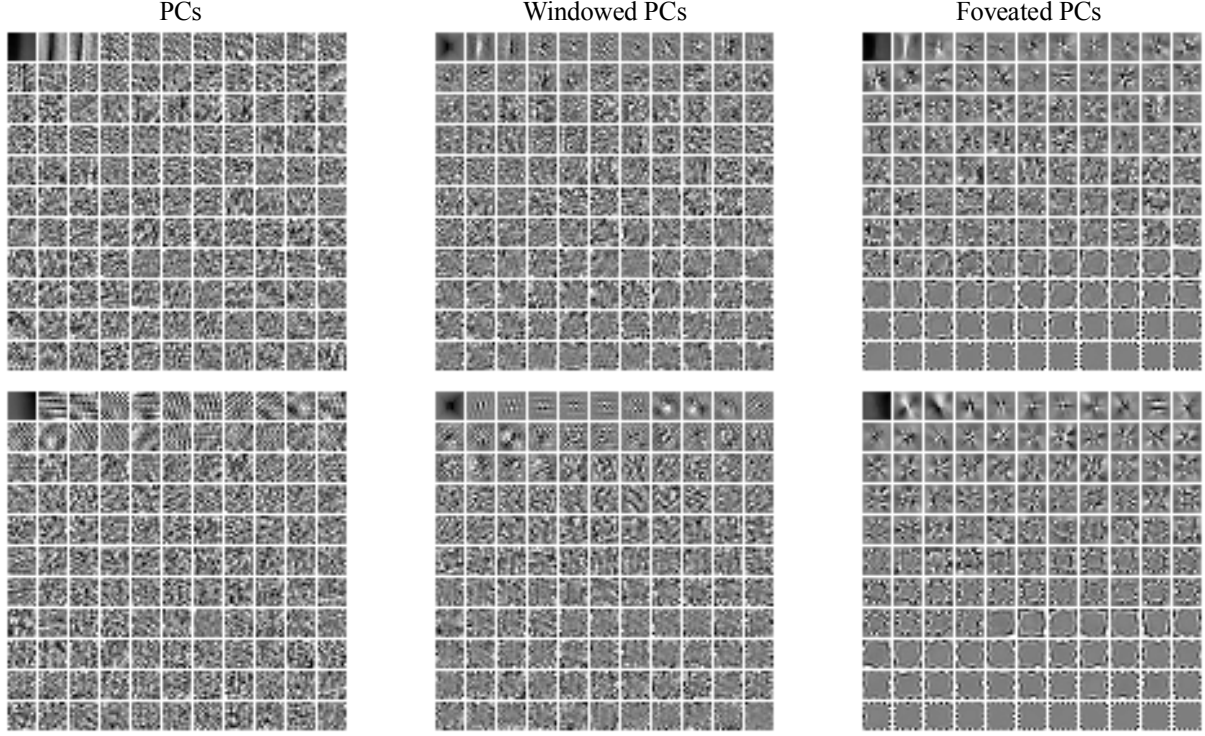
Fig. 1. Examples of the 121 principal components of $11{\times}11$ patches computed from distance-weighted clusters of patches (left), windowed patches (center), or foveated patches (right) (see Section 3) around two distinct centroids (one for the top row and another for the bottom row of subfigures). For each cluster, the distances and weights have been computed from the respective Euclidean (3), windowed (4), or foveated distances (5).

3.1. Analysis

For any patch $\mathbf{z}_j$ in the cluster, we perform analysis (projection on the eigenspaces), as

$$\varphi_{\bar{x},j} = U_{\bar{x}}^T F \overrightarrow{\mathbf{z}_j} = U_{\bar{x}}^T \overrightarrow{\mathbf{z}_j^{\text{FOV}}}, \qquad (8)$$

where F is the matrix representation of the foveation operator, i.e. each row of F is a vectorized PSF $v_u^{\rho,\theta}$, and $\varphi_{\bar{x},j}$ is the spectrum of the foveated patch $\mathbf{z}_j^{\text{FOV}}$ with respect to the foveated PCs.

3.2. Synthesis

Given $\varphi_{\bar{x},j}$, the patch $\mathbf{z}_j$ is reconstructed as

$$\widehat{\mathbf{z}}_{\bar{x},j} = F^{-1} U_{\bar{x}} \varphi_{\bar{x},j}, \qquad (9)$$

where F^{-1} is the (pseudo)inverse of F. Partial or filtered reconstruction is obtained by shrinking, suppressing, or otherwise modifying desired portions of the spectrum $\varphi_{\bar{x},j}$.

3.3. Illustrations

Fig. 1 gives an insight of the effect of foveation by illustrating the PCs computed from clusters of patches, of windowed patches, and of foveated patches. The procedure for computing PCs and windowed PCs is analogous to that for foveated PCs: the employed distances are (3), (4), or (5) (θ=0, ρ=3.5), respectively. Observe how the foveated PCs separate between internal and peripheral features, the former characterizing the first PCs and the latter the last PCs.

4. EXPERIMENTS

To compare the effectiveness of the different PC types, we devise a simple test through the denoising of natural test images corrupted by additive Gaussian noise. The algorithm estimates each patch of the cluster as in (9), but retaining only the first 3 PCs, i.e. all but the first 3 coefficients in the spectrum $\varphi_{\bar{x},j}$ are set to zero. In practice, only the first three columns of $U_{\bar{x}}$ are taken into account in (8) and (9). Overlapping estimates for the same patch are aggregated in a weighted average with weights proportional to $w(\bar{x}, x_j) \, \mathbf{k} \left\| U_{\bar{x}}^T F \right\|_2^{-2}$. These weights are rather natural, and include both the similarity of the patch to the cluster centroid and the noise variance brought by each eigenvector. In the case of windowing (4), F is replaced by a diagonal matrix with diagonal equal to $\sqrt{\mathbf{k}}$; in case of the distance (3), $\mathbf{k}$ is replaced by a constant and F by the identity matrix. The naïve (linear and non-adaptive) shrinkage makes it easier to appreciate the differences due to the (nonlinear and adaptive) construction of the PCs.

Fig. 2 presents the results obtained for two images (σ=25). Due to space limitation, we include here only the reconstruction from windowed and from foveated PCs, noting that the reconstruction from PCs computed from (3) was inferior to both. A close-up inspection reveals that the reconstruction from foveated PCs features more details and better contrast. It is also much less affected by the high-frequency striped ar-

noise-free	noisy (σ=25)	reconstruction from largest 3 windowed PCs	reconstruction from largest 3 foveated PCs

	PSNR = 20.14 dB	PSNR = 27.83 dB	PSNR = 28.29 dB
	PSNR = 20.14 dB	PSNR = 29.11 dB	PSNR = 29.40 dB

Fig. 2. Patch-based reconstruction of natural images from the 3 largest principal components learned from clusters of noisy patches. The reconstruction quality is indicative of how well the PCs are representitive of the patch content.

tefacts typical of dictionaries learned from noisy data. Eigenvectors responsible for these artefacts can be spotted in both the left and the center subfigure on the second row of Fig. 1. The PSNR results are consistent with these observations.

5. CONCLUSIONS

We have extended the nonlocal foveated self-similarity to the clustering of patches, thus introducing the foveated PCs as the eigenvectors of the distance-weighted second-moment matrices of clusters composed of foveated patches. The first few foveated PCs are shown to be more representative of the patch content than either the traditional PCs or windowed PCs, learned from respective clusters of non-foveated patches.

The analysis operates on foveated patches, yet it leads to better decorrelation of the non-foveated content, with the foveation operator F effectively acting as a preconditioner for building the cluster and learning the dictionary.

Ongoing work looks at similarly constructed foveated independent components, as well as at the adaptive selection of the most significant foveated components.

6. REFERENCES

[1] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, 2006.

[2] C.-A. Deledalle, J. Salmon, A. Dalalyan *et al.*, "Image denoising with patch based PCA: local versus global," in *Proc. British Mach. Vision Conf.*, 2011.

[3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "BM3D image denoising with shape-adaptive principal component analysis," in *SPARS'09 - Signal Processing with Adaptive Sparse Structured Representations*, 2009.

[4] P. Chatterjee and P. Milanfar, "Patch-based near-optimal image denoising," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1635–1649, 2012.

[5] A. Foi and G. Boracchi, "Foveated self-similarity in nonlocal image filtering," in *SPIE Human Vision and Electronic Imaging XVII*, vol. 8291. SPIE, 2012.

[6] B. A. Wandell, *Foundations of Vision*. Sinauer Associates, Inc., 1995.

[7] A. Foi and G. Boracchi, "Anisotropic foveated self-similarity," in *Proc. SPARS 2013, Signal Processing with Adaptive Sparse Structured Representations*, July 8-11 2013.

[8] ——, "Anisotropically foveated nonlocal image denoising," in *Proc. IEEE Int. Conf. Image Process. (ICIP 2013)*, September 15-18 2013, pp. 464–468.

[9] 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.

[10] H.-P. Kriegel, P. Kröger, E. Schubert, and A. Zimek, "A general framework for increasing the robustness of PCA-based correlation clustering algorithms," in *Sci. Stat. Database Manag.* Springer, 2008, pp. 418–435.