Appropriate kernels for Divisive Normalization explained by Wilson-Cowan equations

J. Malo*[1] and M. Bertalmio[2]
[1]Image Processing Lab. Parc Científic, Universitat de València, Spain
[2]Dept. Tecnol. Inf. Comunic., Universitat Pompeu Fabra, Barcelona, Spain

Empirically tuning the Divisive Normalization. Cascades of Linear+NonLinear Divisive Normalization transforms [1] can be easily tuned using the derivatives introduced in [2] to reproduce the perception of image distortion in naturalistic environments. Previous brute-force explorations [3] suggested that spatial interactions in divisively normalized wavelets are more relevant to reproduce subjective opinion than scale and orientation interactions. Optimization of such spatial-only kernels confirms this [2]. In this intraband-only Divisive Normalization the vector of V1-like activations, $\boldsymbol{x}$, depends on the energy of linear wavelet responses, $\boldsymbol{e}$, dimension-wise normalized by a sum of neighbor energies,

$$\boldsymbol{x} = \frac{\boldsymbol{e}}{\boldsymbol{b} + H^{\boldsymbol{p}} \cdot \boldsymbol{e}} = \mathbb{D}^{-1}_{(\boldsymbol{b}+H^{\boldsymbol{p}}\cdot\boldsymbol{e})} \cdot \boldsymbol{e} \tag{1}$$

where the kernel $H^{\boldsymbol{p}}$ only considers the departure in spatial position, $\Delta\boldsymbol{p}$, between sensors of the same subband.

Obvious limitations of intraband kernels. Despite successful optimization over large naturalistic image quality databases [2], some basic effects with artificial stimuli may be poorly reproduced [4]: while the model explains cross-orientation and cross-scale masking for low frequency tests seen on high frequency backgrounds it is not the case the other way around. To fix this, a more balanced interaction between subbands in the denominator of Eq. 1 is required, *which cannot be introduced in intraband-only kernels.*

Solution goes beyond Watson & Solomon kernels. The first guess to fix the imbalance is substituting the spatial-only kernel $H^{\boldsymbol{p}}$ in Eq. 1 by more general kernels, as the one proposed by Watson & Solomon, $H^{ws} = H^{\boldsymbol{p}} \odot H^{f} \odot H^{\phi}$, that not only depends on departures in position, $\boldsymbol{p}$, but also in frequency, f, and in orientation ϕ [5]. However, it turns out that Gaussian H^{ws} may not provide the appropriate balance either: low frequency backgrounds may still have too much energy and bias the result for high frequency tests. This may be fixed ad-hoc [4] by *left* and *right* multiplication of the Watson & Solomon kernel with extra diagonal matrices:

$$H = \mathbb{D}_{\boldsymbol{l}} \cdot H^{ws} \cdot \mathbb{D}_{\boldsymbol{r}} \tag{2}$$

While $\mathbb{D}_{\boldsymbol{r}}$, pre-weights the subbands of $\boldsymbol{e}$ before computing the interaction, $\mathbb{D}_{\boldsymbol{l}}$, tunes the relative weight of the masking for each sensor. Additionally to the changes in H to account for the artificial stimuli, the fine-tuned model in [4] included an extra constant to keep the output dynamic range as in the simpler model of Eq. 1, just to keep the previous performance for naturalistic stimuli. Summarizing, the fine-tuned model, $\boldsymbol{x} = \mathbb{D}_{\boldsymbol{k}} \cdot \mathbb{D}^{-1}_{(\boldsymbol{b}+H\cdot\boldsymbol{e})} \cdot \boldsymbol{e}$, requires a specific structure in H, i.e Eq. 2, and vectors $\boldsymbol{l}$ and $\boldsymbol{r}$ of high-pass nature [4].

The question is: *where the structure in Eq. 2 comes from?.*

Explanation: wiring in Wilson-Cowan. The Divisive Normalization model [1] and the Wilson-Cowan model [6] are alternative formulations of the interaction of neural populations. Despite the fact that both models have been shown to have similar information maximization properties [3, 7], and provide similar descriptions of pattern discrimination [8], no direct correspondence has been established between them yet.

In the Wilson-Cowan model [6] the variation of the activation vector, $\dot{\boldsymbol{x}}$, increases with the energy of the input, $\boldsymbol{e}$, but, for each sensor, this variation is also moderated by its own activity and by a linear combination of the activities of the neighbor sensors,

$$\dot{\boldsymbol{x}} = \boldsymbol{e} - \mathbb{D}_{\boldsymbol{\alpha}} \cdot \boldsymbol{x} - \boldsymbol{W} \cdot f(\boldsymbol{x}) \tag{3}$$

where $f(\boldsymbol{x})$ is a dimension-wise sigmoid, and $\boldsymbol{W}$ is the matrix that describes the damping factor between sensors.

In this work we assume that the Divisive Normalization regime is the stationary solution of the dynamic system defined by the Wilson-Cowan equations. We derive the analytical relation between the parameters of both models[1],

$$\begin{aligned} \boldsymbol{b} &\propto \boldsymbol{\alpha} \\ H &\propto \mathbb{D}^{-1}_{\boldsymbol{x}} \cdot \boldsymbol{W} \cdot \mathbb{D}^{-1}_{\boldsymbol{b}} \end{aligned} \tag{4}$$

which is the structure that had to be introduced ad-hoc in Eq. 2. The high pass nature of $\boldsymbol{x}^{-1}$ and $\boldsymbol{b}^{-1}$ explains why the low frequencies in $\boldsymbol{e}$ had to be attenuated by $\boldsymbol{r}$ and $\boldsymbol{l}$.

Discussion. This relation between models has a range of consequences. First, assuming fixed (hard-wired) interaction between the sensors in the Wilson-Cowan model, Eq. 4 implies that the required kernel in Divisive Normalization, H, not only inherits the wiring in $\boldsymbol{W}$, but it also should be signal-dependent. Second, functional forms depending on *proximity* (as in the Watson-Solomon kernel H^{ws}) seem sensible choices for wiring in $\boldsymbol{W}$, which would justify the hand-crafted trick in Eq. 2. Last, but more importantly, Eq. 4 implies that the variety of dynamic analysis already done for Wilson-Cowan systems [9] can also be applied to the wide range of phenomena described by Divisive Normalization.

[1] M. Carandini and D. Heeger. Normalization as a canonical neural computation. *Nature Reviews. Neurosci.*, 13(1):51–62, 2012.

[2] M. Martinez-Garcia, P. Cyriac, T. Batard, M. Bertalmio, and J. Malo. Derivatives and inverse of cascaded linear+noonlinear neural models. *Preprint in arXiv:1711.00526 Quant. Biol.*, 2017.

[3] J. Malo and V. Laparra. Psychophysically tuned divisive normalization approximately factorizes the PDF of natural images. *Neural Computation*, 22(12):3179–3206, 2010.

[4] M. Martinez-Garcia, M. Bertalmio, and J. Malo. In praise of artifice reloaded: Caution with subjective image quality databases. *Preprint in arXiv:1801.09632 Quant. Biol.*, 2018.

[5] A.B. Watson and J.A. Solomon. A model of visual contrast gain control and pattern masking. *JOSA A*, 14:2379–2391, 1997.

[6] H.R. Wilson and J.D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13:55–80, 1973.

[7] M. Bertalmío. From image processing to computational neuroscience: a neural model based on histogram equalization. *Front. Comput. Neurosci.*, 8:71, 2014.

[8] M. Bertalmio, P. Cyriac, T. Batard, M. Martinez-Garcia, and J. Malo. The wilson-cowan model describes contrast response and subjective distortion. *VSS - J. Vision*, 17(10):657–657, 2017.

[9] A. Destexhe and T. Sejnowski. The Wilson-Cowan model 36 years later. *Biol. Cybern.*, 101(1):1–2, 2009.

*Correspondence: jesus.malo@uv.es

[1]Supplementary material (derivation of equivalence) available at: http://isp.uv.es/docs/MODVIS18DNWC.zip