bioRxiv preprint doi: https://doi.org/10.1101/2022.09.17.506431; this version posted September 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Divisive Normalization Circuits Faithfully Represent Auditory and Visual Sensory Stimuli

Aurel A. Lazar[1], Tingkai Liu[1], and Yiyin Zhou[1]

[1]Department of Electrical Engineering, Columbia University
[1]*The authors are listed in alphabetical order*

September 16, 2022

Abstract

Divisive normalization is a canonical neural computation employed for sensory adaptation in vision, olfaction and attention modulation. While Divisive Normalization has been proposed to be an efficient coding algorithm, it remains unclear whether such transformation results in information loss in a dynamic setting. Leveraging a previously proposed general mathematical framework called the Divisive Normalization Processor (DNP), we first show that the DNP circuit describes a wide class of neural circuits as well as phenomenological models including the Linear-Nonlinear cascade model. We then demonstrate both theoretically and computationally that the DNP is an invertible operator that could faithfully represent input information given sufficient output samples.

Executive Summary

Overall Message

Given a *sufficient* number of samples, the Divisive Normalization Processor operating on audio scenes and visual fields is invertible.

Problem Setting

Divisive normalization underlies many models of neural circuits arising in auditory and visual systems. Due to the highly nonlinear nature of the divisive normalization operation in a number of applications such as contrast gain control, there has been a lack of quantitative characterization of the information content of the signals at the output of these circuits. In this paper, we show that, stimuli processed by a class of circuits that we call divisive normalization processors can be faithfully recovered from the outputs of the

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.17.506431; this version posted September 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

circuit, given a sufficient number of samples. We demonstrate this for auditory stimuli (Figure 1A) and visual stimuli (Figure 1B).

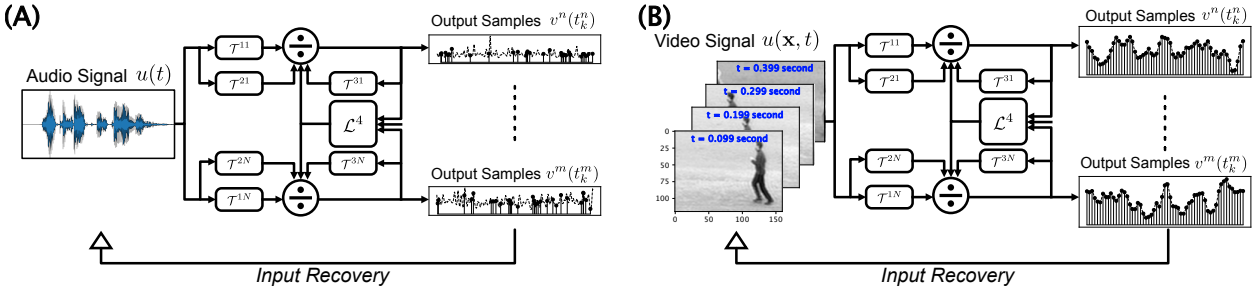

Figure 1: (A) Divisive Normalization Processor (DNP) encoding and input recovery of auditory stimuli. Note that the output signals are asynchronously sampled. (B) DNP encoding and input recovery of video stimuli.

Context/Relevance to the Published Literature

- DNP processing underlies many existing models of auditory systems, e.g., [1].
- For a model of the inner ear with gammatone filterbank and synthetically generated Volterra kernels, audio stimuli can be faithfully recovered from the outputs of the DNP.
- DNP processing underlies many existing models of visual circuits, for example, [2] and [3].

Formal Modeling of DNPs

1. We present the modeling of auditory and visual stimuli.
2. We introduce feedforward and feedback DNPs as depicted in Figure 2, and formally describe the I/O of the DNPs.

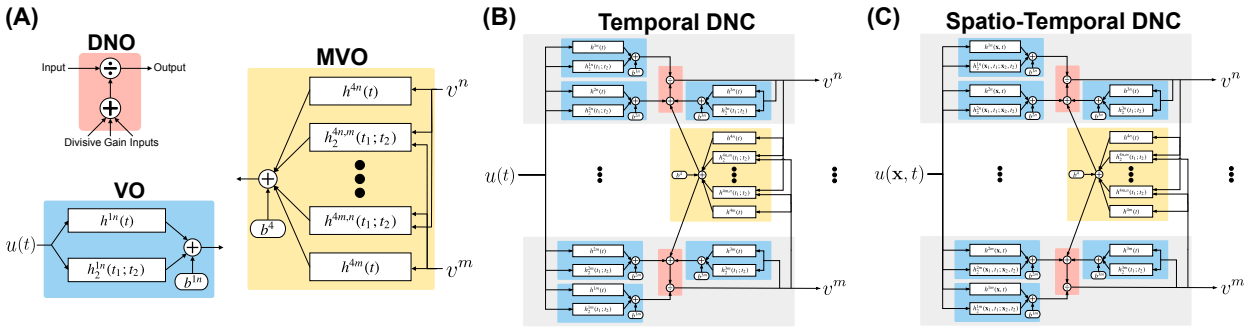

Figure 2: Divisive Normalization Processors. (A) Components of DNPs. (B) DNP for temporal input signal $u(t)$. (C) DNP for spatio-temporal input signal $u(\mathbf{x},t)$.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.17.506431; this version posted September 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Formulating the Recovery of Input Stimuli and Evaluation Criteria

Formulating the Recovery of Input Stimuli

Given the input/output relationship of the DNP as

$$v^n(t) = \frac{\mathcal{T}^{1n}[u](t)}{\mathcal{T}^{2n}[u](t) + \mathcal{T}^{3n}[v^n](t) + \mathcal{L}^4[\mathbf{v}](t)} \tag{1}$$

and irregularly sampled output values $v^n(t_k), n = 1, \ldots, N, k = 1, \ldots, N_K$, the recovery problem is formulated as finding u given the samples $v^n(t_k)$ with known operators $\mathcal{T}^{1n}, \mathcal{T}^{2n}, \mathcal{T}^{3n}, \mathcal{L}^4$.

Evaluation Criteria

Given the recovered input signal $\tilde{u}$, we evaluate the recovery via Signal-to-Noise Ratio (SNR) as well as Peak Signal-to-Noise Ratio (PSNR):

$$SNR(\hat{u}, u) = 10 \cdot \log_{10} \frac{\int_{\mathbb{D}} |u|^2(t)dt}{\int_{\mathbb{D}} |u - \hat{u}|^2(t)dt}, \quad PSNR(\hat{u}, u) = 10 \cdot \log_{10} \frac{\|u\|_\infty^2}{\int_{\mathbb{D}} |u - \hat{u}|^2(t)dt},$$

where $u = u(t)$ is the temporal stimulus, $\hat{u} = \hat{u}(t)$ is the recovered stimulus. For spatio-temporal stimuli, $u = u(\mathbf{x}, t)$ is the stimulus and $\hat{u} = \hat{u}(\mathbf{x}, t)$ is the recovery. $\mathbb{D}$ defines the domain of the stimuli.

Perfect Recovery Algorithms

Stimuli Processed by DNPs with First Order Kernels Are Recovered via Regularized Regression

- Sampling of the DNP output can be asynchronous (this is a processor with continuous input and continuous output)
- The reconstructed stimulus is the solution of the *convex* optimization problem

$$\min_{\mathbf{u} \in \mathbb{C}^{2L+1}, \mathbf{v}^n \in \mathbb{C}^{2L^o+1}, n=1,\ldots,N, \boldsymbol{\varepsilon} \in \mathbb{R}^{N \cdot N_K}} \|\boldsymbol{\varepsilon}\|_2 + \lambda \cdot (\|\mathbf{u}\|_2 + \sum_{n=1}^{N} \|\mathbf{v}^n\|_2)$$

$$\text{subject to} \quad \left[\begin{array}{c} (v_k^n \mathbf{h}_k^{2n} - \mathbf{h}_k^{1n})^T \mathbf{u} + (v_k^n \mathbf{h}^{3n})^T \mathcal{P}_{\mathcal{H}_1^o} \mathbf{v}^n + v_k^n \sum_{m=1}^{N} (\mathbf{h}_k^{4m})^T \mathcal{P}_{\mathcal{H}_1^o} \mathbf{v}^m \\ = b^{n1} - v_k^n (b^{n2} + b^{n3} + b^4) + \varepsilon_k^n \end{array} \right]$$

$$\mathbf{1}^T \boldsymbol{\varepsilon} = 0$$

where the expression in $[\cdot]$ brackets is the I/O constraint of the DNP with ε_k^n denoting the error of measurement at the k-th sample of the n-th neuron.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.17.506431; this version posted September 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

- Without feedback, the total number of samples required is $2L+1$ (same as the order of the input signal), with feedback the total number of samples required is $2L+1+N(2L^o+1)$ (proportional to the number of channels). The extra samples are needed to recover the outputs from the samples according to the Shannon-Nyquist sampling theorem.

Stimuli Processed by DNPs w/ Second Order Kernels Are Recovered via Rank-Minimization

- With $\mathbf{U} = \mathbf{u}\mathbf{u}^\dagger, \mathbf{V}^{nm} = \mathcal{P}_{\mathcal{H}_1^o}\mathbf{v}^n(\mathcal{P}_{\mathcal{H}_1^o}\mathbf{v}^m)^\dagger$, where † denotes the conjugate transpose, and by observing that $\mathbf{U}$ and $\mathbf{V}^{nm}$ are rank 1, we formulate the recovery problem as a rank-minimization problem with a nuclear-norm *convex* relaxation:

$$\min_{\mathbf{u},\mathbf{v}^n,\mathbf{U},\mathbf{V}^n,\boldsymbol{\varepsilon}} \|\mathbf{U}\|_* + \sum_n \|\mathbf{V}^n\|_* + \lambda_1 \cdot \|\boldsymbol{\varepsilon}\| + \lambda_2 \cdot (\|\mathbf{u}\| + \sum_n \|\mathbf{v}^n\|)$$

$$\text{s.t.} \quad \begin{bmatrix} (v_k^n \mathbf{h}_k^{2n} - \mathbf{h}_k^{1n})^T \mathbf{u} + (v_k^n \mathbf{h}^{3n})^T \mathcal{P}_{\mathcal{H}_1^o}\mathbf{v}^n + v_k^n \sum_{m=1}^N (\mathbf{h}_k^{4m})^T \mathcal{P}_{\mathcal{H}_1^o}\mathbf{v}^n \\ + \mathrm{Tr}((v_k^n \mathbf{H}_{2k}^{2n} - \mathbf{H}_{2k}^{1n})^T \mathbf{U}) + v_k^n \sum_{l=1}^N \sum_{m=1}^N \mathrm{Tr}((\mathbf{H}^{4lm})^T \mathbf{V}^{lm}) \end{bmatrix}$$
$$= b^{1n} - v_k^n (b^{2n} + b^{3n} + b^4) + \varepsilon_k^n, \ \forall n, k$$
$$\mathbf{U} \succeq 0$$
$$\mathbf{V}^n \succeq 0 \ \forall n$$
$$\mathbf{1}^T \boldsymbol{\varepsilon} = 0$$

  where the first constraint above is the I/O of DNP with first and second order kernels, ε_k^n is the error of the measurement at the k-th sample of the n-th neuron, $\|\cdot\|_*$ denotes the nuclear norm operator, and $\lambda_1, \lambda_2 \geq 0$.

- Given a sufficient number of samples, the recovery is perfect and the recovered 2nd order coefficient is rank-1.

We use simple temporal stimuli to characterize how well the stimulus is recovered with an increasing number of samples. Perfect recovery can always be achieved when the number of samples is above a threshold (Figure 3).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.17.506431; this version posted September 18, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

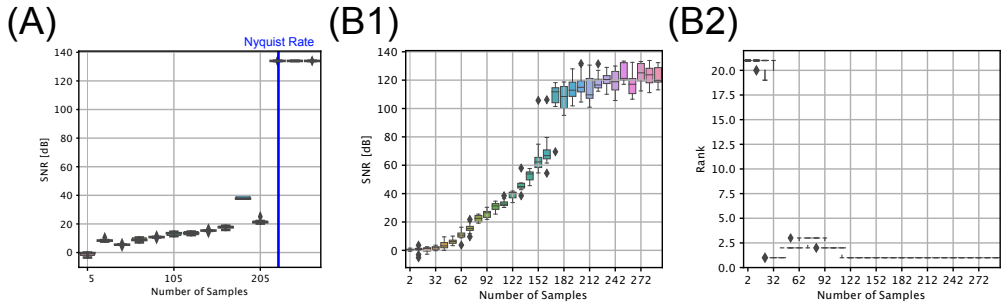

Figure 3: Input recovery for DNP circuit with 1st (A) and 2nd (B1,B2) order kernels. (A,B1) Given a sufficient number of samples, the input recovery for DNP circuit reaches Signal-to-Noise ratio above 133 dB, indicating perfect recovery. (B2) For DNP circuit with 2nd order kernels, perfect input recovery correponds to when the recovered 2nd order input coefficient **U** has rank 1.

Acknowledgments

The research reported here was supported by AFOSR under grant #FA9550-16-1-0410, DARPA under contract #HR0011-19-9-0035 and NSF under grant #2024607.

References

[1] O. Schwartz and E. P. Simoncelli, "Natural sound statistics and divisive normalization in the auditory system," in *Advances in Neural Information Processing Systems*, 2001.

[2] M. Carandini and D. J. Heeger, "Normalization as a canonical neural computation," *Nature Reviews Neuroscience*, vol. 13, pp. 51–62, 2012.

[3] J. H. Van Hateren and H. P. Snippe, "Information theoretical evaluation of parametric models of gain control in blowfly photoreceptor cells," *Vision Research*, vol. 41, no. 14, pp. 1851–1865, 2001.