

Narrowing the gap between analytically convenient and naturalistic stimuli

Jonathan D. Victor

Feil Family Brain and Mind Research Institute
Weill Cornell Medicine
1300 York Avenue
New York, NY 10065

Voice: 212 746 2343
Email: jdvicto@med.cornell.edu

Running head: Narrowing the analytic-naturalistic gap

Keywords: naturalistic stimuli; systems analysis, efficient coding, visual textures, generative AI

Precis

Two approaches to characterizing a visual process – use of simple stimuli vs. naturalistic ones – are often considered to be philosophically divergent. This divergence may be more apparent than real, and will be bridged as we achieve a greater understanding of the statistics of natural scenes.

Abstract

The gulf between mathematically-motivated stimuli -- such as sine waves and noise ensembles -- and images drawn from natural scenes may not be as wide as it appears. For the former, the mathematical properties that support their utility as analytic probes are explicit. For naturalistic stimuli, this is not the case, but they are created by physical processes whose outcomes -- at least in principle -- heavily constrain their statistics. Images generated computationally occupy a middle ground: their mathematical properties remain implicit in the rules that generate them, but these rules are simpler than the processes that generate natural images. It is suggested that a greater understanding of the statistics of natural images, while challenging, will enhance their utility as experimental tools and thereby lead to new insights into the design principles of sensory systems.

Perspective

A main theme of this Special Issue is the choice between naturalistic and artificial stimuli, and the relative merits of these stimulus classes is perhaps the most persistently and passionately debated aspects of experimental design (Felsen & Dan, 2005; Rust & Movshon, 2005; Sharpee, 2013). Discussions on this point often conclude with an agreement to disagree, citing the purpose of the experiment and perhaps the experimenter's philosophy. Here, I suggest greater common ground: that there is fundamental agreement about a core scientific goal, and the naturalistic vs. artificial debate reflects an incomplete state of both conceptual and biological understanding, leading to uncertainty as to how the goal is best achieved, and a gap that appears wider than it really is. The common goal is a simple one: the need to characterize a system's behavior across a substantial range of its inputs, and to do so using stimuli that drive the normal physiological mechanisms. In attempting to support this claim, I hope that this perspective will provide a common framework for thinking about many stimulus classes, including those that are naturally parameterized in the spatial domain, those that are naturally parameterized in the Fourier domain, as well as noise stimuli, artificial and naturalistic textures, and natural scenes. I also hope that it will suggest directions for new theoretical investigations.

Under some circumstances, stimulus choice is indeed driven by the purpose of the experiment. If the immediate goal is to measure the response to a particular set of naturalistic stimuli, then that stimulus set is the obvious and logical choice. But if the goal is to understand the neural computations underlying how a system responds in general, the decision is less clear. There the default choice shifts to artificial stimuli: they support stringent tests of a model, including insights into its limitations. Fundamentally, this is often because artificial stimuli are “analytically convenient:” they constitute a library of simple stimuli that, because of their mathematical properties, can be used to build up more complex ones and to test models.

As it is nearly tautologous that the first goal is well-served by selecting natural stimuli, I will attempt to support the claim that the rationales for stimulus choice are convergent by taking the viewpoint of a computational modeler. Here, experimental design requires selection of stimuli from a space with an enormous number of degrees of freedom: in the case of vision, how light intensity varies as a function of two (or three) spatial dimensions, one temporal dimension, and perhaps wavelength. The first step in making this problem tractable is to assume that many of these degrees of freedom are irrelevant: for example, that only a restricted region of space (the receptive field!), and only a particular amount of its immediate past history is important. With such assumptions, along with choosing a spatial resolution and a temporal resolution, the number of degrees of freedom is large but manageable.

However, this restricted space of possible stimuli is far too large to explore by brute force. The standard next step is to make an assumption about the stimulus-response relationship. For neurons in early visual processing, one way to proceed is to assume that linearity is a good approximation: that addition (superposition) of stimuli results in addition of the responses. A related but weaker assumption is that the output results from an additive interaction followed by a pointwise nonlinearity (i.e., one that applies the same transformation at every point in time). Given an assumption of this kind -- and we will consider alternatives below -- it suffices to probe the input space one dimension at a time, and only a single stimulus on each dimension is necessary (or a few stimuli, if one allows for a pointwise nonlinearity). That is, determining the response of the system to each of the basis vectors in the stimulus space suffices to fully characterize the system. The problem of characterizing responses to the entire stimulus space reduces to the problem of characterizing responses along a set of axes; the number of axes is equal to the number of dimensions in the space. In principle, the number of such dimensions is infinite, but, with assumptions about temporal and spatial resolution limits, the number is merely very large.

If the system is linear for one basis set, it is linear for all other basis sets. So, an important choice remains: which axes? It makes intuitive sense that an efficient approach is to choose the axes – i.e., the spatiotemporal patterns used to probe the system – so that they are as different as possible from each other, to minimize the redundancy between measurements. Geometrically, this corresponds to axes that are orthogonal, as a correlation between spatiotemporal patterns corresponds to a dot-product between their representations as vectors. But even within the realm of orthogonal stimulus sets, there are many possibilities: orthogonality only determines the axes up to rotation. One choice is a stimulus set consisting of spots of light at a chosen spatial resolution, each presented for a duration corresponding to the chosen temporal resolution (“the spacetime” basis set). A second choice is the “Fourier basis”, a stimulus set consisting of sinusoids in space, modulated sinusoidally in time. There are also a plethora of intermediate choices, including gratings presented as flashes, Gabor functions(Daugman, 1985; Gabor, 1946), Hermite functions, and their generalizations(J.D. Victor & Knight, 2003). Each of these choices amount to choosing a different set of orthogonal (or, in the case of Gabor functions, approximately orthogonal) coordinate axes.

Two general considerations that impact on this choice are illustrated by comparing the spacetime basis with the Fourier basis.

One consideration flows from the approximate nature of linearity. Linearity is guaranteed to be violated for stimuli whose magnitude is sufficiently large; conversely, this approximation is more likely to be useful if stimulus magnitude is limited. On the other hand, the realities of biological and measurement noise mean that one wants stimuli that produce a response that is not close to noise levels. In characterizing a system with a spacetime basis, the spot size and duration need to be small, to match the temporal and spatial resolution of the system. But to achieve a non-negligible response with such a stimulus, the pulse amplitude may need to be large. An abrupt transition from background to a small,

bright flash may be the worst-case-scenario in terms of what is needed to give the assumption of linearity a chance to be effective.

In contrast, the Fourier basis calls for stimuli that are smoothly modulated in space and time. There is of course a tradeoff: one now has to decide how to sample in the Fourier domain. For many physical systems, characteristics depend smoothly on frequency, so a relatively sparse sampling will suffice. Note also that the strategy of attempting to avoid saturations by seeking a set of stimuli that produce large responses for a given input size amounts to seeking to match the stimuli to the receptive field profile. Often, this corresponds to choices that are intermediate between the spatiotemporal domain and the Fourier domain (Pollen, Nagler, Daugman, Kronauer, & Cavanagh, 1984).

A second, and very different, consideration is that it is often helpful and reassuring to have a direct path from the experimental data to some understanding of the system under study. Both the spacetime strategy and the Fourier strategy have advantages here. The average response to impulses at each spatial location is a direct characterization of the system, and, under the hypothesis of linearity, the superposition of impulse responses suffice to predict the response to any stimulus. Its Fourier transform (i.e., the transfer function), which is directly measured by the responses to sinusoidal stimuli, has convenient analytic properties: the transfer function of subsystems in parallel, in series, or in feedback loops combine by addition, multiplication, and division. Another way that such responses to such stimulus sets can provide insight is that the need for nonlinear components is immediately evident in a frequency-domain approach: a strictly linear system will only produce outputs at the stimulation frequency, and not at its harmonics.

While these considerations may be useful conceptually, their practical utility fades when one is more than one synapse away from the sensory periphery. Even if approximate linearity holds, neurons may be narrowly tuned: geometrically, this means that there are few directions in stimulus space that

lead to responses at all. Nonlinearity only compounds the problem: a nonlinear neuron may respond very differently to a combination of stimulus components, than is predicted by summing the response to each of the components presented in isolation, so characterization of the response along several axes in stimulus space do not characterize the behavior for off-axis points.

Approaches based on noise stimuli – spatiotemporal Gaussian noise, spatiotemporal binary noise, and related strategies – may be viewed as attempts to address this problem, but they typically fall short. Each of these stimulus sets sample a region around the origin, without restriction to a preferred set of axes. However, while these approaches have proven effective for characterizing neurons in retina and V1 that have relatively simple nonlinearities and are broadly tuned (Meister & Berry, 1999; Reid, Victor, & Shapley, 1997), they miss critical aspects of neural selectivity in later visual areas, such as V4 (Gallant, Braun, & Van Essen, 1993; Pasupathy & Connor, 2001) and beyond. The reason is straightforward: while a noise ensemble broadly samples the space of possible images, the sampling of any particular image is so infrequent so as to be nonexistent in practical terms. For such neurons – and more generally, for neurons that are highly nonlinear or narrowly tuned -- naturalistic stimuli can thus be viewed as providing a shortcut to finding the regions of stimulus space in which a visual neuron will respond, and avoiding the much larger volume of stimulus space that is irrelevant.

These considerations concerning choice of a stimulus set apply equally well – if not more so – when the experimenter is recording from large populations of neurons, or describing the response in more abstract terms, such as a response manifold (Langdon, Genkin, & Engel, 2023). The output variable is now more complex or more abstract than a single time series of spikes, but the central issue is the same: how should the set of stimuli be chosen, so that the full range of neural behavior can be observed?

I expect that the comments up to this point are non-controversial, perhaps even obvious. But now I would like to take a more speculative viewpoint: that naturalistic stimuli will ultimately turn out to

have much more in common with “constructed” stimulus sets than we currently imagine -- it’s just that we don’t yet have the insights to recognize this.

The first claim is that we may be far from what is achievable in terms of a mathematical characterization of naturalistic stimuli. We know that naturalistic stimuli are a miniscule subset of all possible stimuli, and we have some insights as to what this subset is, but it appears quite incomplete. For example, naturalistic stimuli have distinctive statistics (e.g. the $1/f^2$ power spectrum(Field, 1987)) – but images created with this power spectrum and no other constraints do not in any way look natural. At least part of the reason is that naturalistic stimuli have other statistical properties not captured by the power spectrum (for example, local image statistics(Hermundstad et al., 2014; Tkacik, Prentice, Victor, & Balasubramanian, 2010) and correlations across frequency bands(Schwartz & Simoncelli, 2001)). It is a difficult mathematical problem to sample randomly from images with statistical constraints on the spectra and higher-order quantities, so it is unclear whether these relatively simple statistical constraints suffice to account for the “natural” subset of all possible stimuli. Intuitively, a large number of degrees of freedom remain, but surprisingly, work by Wakita and Motoyoshi (Wakita & Motoyoshi, 2026) has shown that a wide range of naturalistic textures can be captured by as few as 16 latent variables. This emphasizes the remaining gap between current mathematical characterizations of natural images, and empirical evidence.

It is of course risky to claim that an insight is achievable, without providing a path to achieve it. So here is a thought experiment to explain my optimism. Imagine being a vision scientist at a time in which the standard mathematical functions were just the polynomials, and one was tasked with describing a movie of an object on a spring. Basic physics yields the differential equation that governs its motion: acceleration, which is proportional to force, is a linear function of position. Since acceleration is the second derivative of position, one could then calculate the motion path numerically.

Now it is also well-known that the solutions to this differential equation are simply the sines and cosines of an appropriate frequency. In the absence of a compact notation for trigonometric functions, one could be forgiven for considering these solutions to be awkward objects, since they are not readily expressed in terms of the familiar elements (i.e., the polynomials), and hence, might be difficult to work with. But importantly, the solutions to the equations of motion have some really nice properties -- periodicity and orthogonality -- and these properties can be demonstrated directly from the motion equation, without ever producing a simple formula for its solutions. This changes how the solutions are viewed: from awkward objects known only numerically to indispensable elements of mathematical analysis.

Two approaches to the study of visual texture provide for a transition from a thought experiment to real experiments. The well-known texture-synthesis procedure of Portilla and Simoncelli (Portilla & Simoncelli, 2000) generate naturalistic images whose statistical properties are implicit in the iterative algorithm that generates them. But even without an explicit characterization of these statistics, they are useful tools in characterizing the computations carried out in early visual areas (Freeman, Ziemba, Heeger, Simoncelli, & Movshon, 2013).

In our laboratory, a set of texture-synthesis algorithms (Victor & Conte, 2012), loosely inspired by the dynamics of a physical system of interacting spins (Pickard, 1980), generates a high-dimensional space of visual textures. For textures in which the interactions have simple forms (pairwise and some higher-order cases), their correlation structure can be described with standard exponentials (Victor & Conte, 2022). But for more complex interactions, no such simple representation appears possible. Nevertheless (Victor & Conte, 2012), these textures have a “maximum entropy” property, which follows directly from the texture generation procedure, i.e., they are as random as possible given the local rules that generate them. The maximum-entropy property allows a normative model to connect perceptual

thresholds measured with these textures to the statistics of natural images (Hermundstad et al., 2014; Tesileanu et al., 2020; Tkacik et al., 2010).

Natural scenes and the images they form are, of course, quite complex: they are driven by multiple kinds of processes that act at multiple scales, and we don't fully understand them. Nevertheless, even in this scenario, insights are possible. As an example: the characteristic $1/f^2$ spectrum of natural scenes can be derived from a simple symmetry argument, along with the heuristic that a natural scene consists of occlusive, discrete objects: when one "steps into" a natural scene, another natural scene results: some objects are now behind the viewer, and distant objects become closer, subtending a larger angle. The $1/f^2$ spectrum follows from the observation that the resulting *statistics* cannot change: it is the only spectrum that has this property. Interestingly, a variation of this argument applies to scenes composed of partially transparent objects that fill space and are viewed at a fixed distance (as in some medical radiographs, such as mammograms), here, the spectrum changes to $1/f^3$ (Burgess, Jacobson, & Judy, 2001)).

In the thought experiment of a movie of an oscillating weight, the key mathematical properties of the motion (e.g., the periodicity and orthogonality of sines and cosines) can be derived from the generative process, without producing explicit solutions. In the texture-synthesis algorithms of Victor & Conte (2012), this holds for the maximum-entropy property, and in the natural-scenes example, it holds for the power spectrum. In the former case, these mathematical properties underlie the utility of these stimulus sets as analytic tools; in the latter, it enables fundamental insights about the design principles of early vision (Field, 1987; Kuang, Poletti, Victor, & Rucci, 2012). The hope is that these examples represent starting points for further progress, rather than bounds. If so, then strategies that focus on the physical processes that generate natural images will prove similarly fruitful, leading directly to identification of their distinctive mathematical properties without a requirement for explicit synthesis.

Interestingly, we are already comfortable with the use of stimulus sets defined by computational generative processes. Textures typically fall into this category: a set of moments or distributional properties of an image ensemble are specified, along with a procedure for creating exemplars. The creation procedure is based on the heuristic of choosing a “random” example of an image drawn from the ensemble. In some cases, the procedure is an explicit sequence of steps: phase randomization (Freeman et al., 2013) or a Markov process (Julesz, 1962; Victor & Conte, 2012) and indeed leads to stimuli that have convenient mathematical properties. In others, the mathematical properties of the outcome are less clear. This includes generative processes that are a one-pass sequence of steps (Skriabine, Shinn, Picard, Harris, & Carandini, 2026), as well as those that are iterative (e.g., the Portilla-Simoncelli (2002) procedure, and later variants (Freeman & Simoncelli, 2011; Lieber, Oleskiw, Palmieri, Simoncelli, & Movshon, 2025)).

More elaborate generative methods, taking their algorithms from machine vision have also proven their worth, in revealing general features of tuning properties in V1 and V4 (Franke et al., 2026), in defining the selectivity of inferotemporal neurons (Ponce et al., 2019), and in identifying causal relationships between activity in inferotemporal cortex and perception (Shahbazi, Ma, Pernus, Scheirer, & Afraz, 2024). These generative processes are distinct from the generative processes that underlie real-world images – and the images that they generate are not natural, but – to human observers – have recognizable, naturalistic features, though are not nameable. The images activate visual neurons deep within the visual system quite well – indeed, demonstrably better than real-world natural stimuli (Bashivan, Kar, & DiCarlo, 2019; Ponce et al., 2019). Moreover, these stimuli are particularly incisive stress tests of network models of V4 neurons: they reveal the ability of models to predict how such stimuli can drive neural activity outside the range that is observed in response to natural images (Bashivan et al., 2019), but also, substantial discrepancies between responses of real V4 neurons

and the models(Ren & Bashivan, 2024). One may even speculate that the quantitative differences revealed by such stimuli are reflections of more fundamental qualitative differences (e.g, “off-manifold” vs. “on-manifold”(Oby et al., 2025) activity patterns), when ensemble responses are analyzed by geometric means(Langdon et al., 2023).

Physical processes are not the same as the generative processes used in machine vision, but also, the visual system does not require that stimuli are strictly natural. Visual systems are well-adapted to our visual world, but they are not so narrowly tuned to “natural” images that they cannot make sense of things that are novel(Schmittwilken, Haverkamp, & Maertens, 2025). The visual system – like many other sensory systems– is shaped by evolution and development to work efficiently for inputs that are natural(Barlow, 1961), but it is enough of a generalist to behave usefully for inputs unlike those that have been previously encountered. The strategies that it uses to achieve this performance are as yet unclear, and uncovering them may eliminate the apparent gap between natural and computationally convenient stimuli.

In sum, I am suggesting that the gap between analytically convenient stimuli and naturalistic ones may not be as wide as it appears. Naturalistic stimuli are created by physical processes whose outcomes, at least in principle, are fully susceptible to a mathematical description. At present, our characterizations of natural stimuli and their statistics – while clearly important for understanding visual processing -- are far from complete. Bridging this gap is a challenge for future work, and even narrowing it is likely to lead to new insights into the design principles of sensory systems.

Acknowledgements

The author deeply appreciates Mary M. Conte's many helpful comments on this manuscript, and would like to acknowledge Alon Amit's provocative comments on Quora concerning "closed form" expressions, especially Quora answer 1236654. This work was supported by NIH NEI EY7977.

References

- Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217-234). Cambridge, MA: MIT Press.
- Bashivan, P., Kar, K., & DiCarlo, J. J. (2019). Neural population control via deep image synthesis. *Science*, 364(6439). doi:10.1126/science.aav9436
- Burgess, A. E., Jacobson, F. L., & Judy, P. F. (2001). Human observer detection experiments with mammograms and power-law noise. *Med Phys*, 28(4), 419-437. doi:10.1118/1.1355308
- Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J Opt Soc Am [A]*, 2(7), 1160-1169.
- Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. *Nat Neurosci*, 8(12), 1643-1646. doi:10.1038/nn1608
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Am [A]*, 4(12), 2379-2394.
- Franke, K., Karantzas, N., Willeke, K., Diamantaki, M., Ramakrishnan, K., Bedel, H. A., . . . Tolias, A. (2026). Dual-feature selectivity enables bidirectional coding in visual cortical neurons. *Elife*. doi:10.7554/eLife.109861.1
- Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nat Neurosci*, 14(9), 1195-1201. doi:10.1038/nn.2889
- Freeman, J., Ziemba, C. M., Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (2013). A functional and perceptual signature of the second visual area in primates. *Nat Neurosci*, 16(7), 974-981. doi:10.1038/nn.3402
- Gabor, D. (1946). Theory of communication. *J. IEEE London*, 93(III), 429-457.

- Gallant, J. L., Braun, J., & Van Essen, D. C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science*, 259(5091), 100-103.
- Hermundstad, A. M., Briguglio, J. J., Conte, M. M., Victor, J. D., Balasubramanian, V., & Tkacik, G. (2014). Variance predicts salience in central sensory processing. *Elife*, 3. doi:10.7554/eLife.03722
- Julesz, B. (1962). Visual pattern discrimination. *IRE Trans Inf Theory*(IT-8), 84-92.
- Kuang, X., Poletti, M., Victor, J. D., & Rucci, M. (2012). Temporal encoding of spatial information during active visual fixation. *Curr Biol*, 22(6), 510-514. doi:10.1016/j.cub.2012.01.050
- Langdon, C., Genkin, M., & Engel, T. A. (2023). A unifying perspective on neural manifolds and circuits for cognition. *Nat Rev Neurosci*, 24(6), 363-377. doi:10.1038/s41583-023-00693-x
- Lieber, J. D., Oleskiw, T. D., Palmieri, L., Simoncelli, E. P., & Movshon, J. A. (2025). Responses of neurons in macaque V4 to object and texture images. *bioRxiv*. doi:10.1101/2024.02.20.581273
- Meister, M., & Berry, M. J., 2nd. (1999). The neural code of the retina. *Neuron*, 22(3), 435-450.
- Oby, E. R., Degenhart, A. D., Grigsby, E. M., Motiwala, A., McClain, N. T., Marino, P. J., . . . Batista, A. P. (2025). Dynamical constraints on neural population activity. *Nat Neurosci*, 28(2), 383-393. doi:10.1038/s41593-024-01845-7
- Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: position-specific tuning for boundary conformation. *J Neurophysiol*, 86(5), 2505-2519.
- Pickard, D. K. (1980). Unilateral Markov fields. *Adv. Appl. Probab.*, 12, 655-671.
- Pollen, D. A., Nagler, M., Daugman, J., Kronauer, R., & Cavanagh, P. (1984). Use of Gabor elementary functions to probe receptive field substructure of posterior inferotemporal neurons in the owl monkey. *Vision Res*, 24(3), 233-241.

Ponce, C. R., Xiao, W., Schade, P. F., Hartmann, T. S., Kreiman, G., & Livingstone, M. S. (2019).

Evolving Images for Visual Neurons Using a Deep Generative Network Reveals Coding

Principles and Neuronal Preferences. *Cell*, 177(4), 999-1009 e1010.

doi:10.1016/j.cell.2019.04.005

Portilla, J., & Simoncelli, E. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *Int'l Journal of Computer Vision*, 40(1), 49-71.

Reid, R. C., Victor, J. D., & Shapley, R. M. (1997). The use of m-sequences in the analysis of visual neurons: Linear receptive field properties. *Visual Neuroscience*, 14(6), 1015-1027. Retrieved from <Go to ISI>://WOS:A1997YG59600002

Ren, Y., & Bashivan, P. (2024). How well do models of visual cortex generalize to out of distribution samples? *PLoS Comput Biol*, 20(5), e1011145. doi:10.1371/journal.pcbi.1011145

Rust, N. C., & Movshon, J. A. (2005). In praise of artifice. *Nat Neurosci*, 8(12), 1647-1650.

Schmittwilken, L., Haverkamp, A. L., & Maertens, M. (2025). Line drawings as a tool to probe edge sensitivity in natural scenes. *J Vis*, 25(14), 22. doi:10.1167/jov.25.14.22

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nat Neurosci*, 4(8), 819-825.

Shahbazi, E., Ma, T., Pernus, M., Scheirer, W., & Afraz, A. (2024). Perceptography unveils the causal contribution of inferior temporal cortex to visual perception. *Nat Commun*, 15(1), 3347. doi:10.1038/s41467-024-47356-8

Sharpee, T. O. (2013). Computational identification of receptive fields. *Annu Rev Neurosci*, 36, 103-120. doi:10.1146/annurev-neuro-062012-170253

Skriabine, S., Shinn, M., Picard, S., Harris, K. D., & Carandini, M. (2026). Mapping the visual cortex with Zebra noise and wavelets. *J Vis*, 26(1), 1. doi:10.1167/jov.26.1.1

- Tesileanu, T., Conte, M. M., Briguglio, J. J., Hermundstad, A. M., Victor, J. D., & Balasubramanian, V. (2020). Efficient coding of natural scene statistics predicts discrimination thresholds for grayscale textures. *Elife*, 9. doi:10.7554/eLife.54347
- Tkacik, G., Prentice, J. S., Victor, J. D., & Balasubramanian, V. (2010). Local statistics in natural scenes predict the saliency of synthetic textures. *Proceedings of the National Academy of Sciences of the United States of America*, 107(42), 18149-18154. doi:10.1073/pnas.0914916107
- Victor, J. D., & Conte, M. M. (2012). Local image statistics: maximum-entropy constructions and perceptual salience. *J Opt Soc Am A Opt Image Sci Vis*, 29(7), 1313-1345. doi:10.1364/JOSAA.29.001313
- Victor, J. D., & Conte, M. M. (2022). Functional recursion of orientation cues in figure-ground separation. *Vision Res*, 197, 108047. doi:10.1016/j.visres.2022.108047
- Victor, J. D., & Knight, B. W. (2003). Simultaneously band and space limited functions in two dimensions, and receptive fields of visual neurons. In E.Kaplan, J. Marsden, & K. R. Sreenivasan (Eds.), *Springer Applied Mathematical Sciences Series* (pp. 375-420): Springer.
- Wakita, S., & Motoyoshi, I. (2026). A compact perceptual space for natural textures emerges from natural image statistics. *Vision Res*, 243, 108795. doi:10.1016/j.visres.2026.108795