Stimulus overview:

The table below lists a breakdown of stimuli signals, along with key luminance, frequency, phase, psychophysical, neurophysiological, and other factors. They are arranged from the simplest excitation of visual processes (a flat field) to the most complex (civilized imagery).

Signal	Luminance	Frequency characteristics	Phase characteristics	Psychophysical aspects	Neurophysiological aspects	Other comments
Flat field 1 *	Single luminance level; Contrast = 0	DC	None	Surround effects; FOV issues known; 'crispening effect'	Overall light adaptation	
Single Gabor 2 *	Contrast via multiple luminances (Michelson or RMS contrast); Surrounding area has constant luminance	Narrowband frequency; Narrowband orientation; single position in Fourier quadrant with Gaussian spread	Coherent phase	Known as most detectable signal in terms of lowest contrast energy; both masking and facilitation	Response of simple cells in V1; localized to cortical area corresponding to retinal position; Position is affected by CMF. Limited to cortical columns corresponding to the orientation	Sines have problems and excite a wider cortical region; simple patches even worse. Our experiment has a Gabor with frequency of 10.8 cy/picture
Field of Gabors with single orientation, single radial frequency 3 *	Same as above; surrounding area of constant luminance can be avoided	Same as above;	Coherent phase or incoherent can be used	Higher detectability due to probability summation and FOV aspects; No facilitation with incoherent phase	Stimulating simple cells throughout the cortex per the extra positions; Still limited to regions corresponding to the orientation	Incoherent phase is used to prevent comparison of mean level surrounding the Gabor
Field of Gabors with variable orientation (preferred= random), but same radial frequency 4 *	Same as above	Ring in the Fourier domain whose radius is the radial frequency.	Same as above, but difference is less due to orientation variability	Sam as above, except little chance of facilitation even with coherent phase	Stimulates entire region of V1 simple cells for the center frequency; all columns of orientation; CMF regions corresponding to the FOV	Same as above
Bandpass	No regions	A blob in the	Incoherent	Orientation	Like above, but	Not used in

filtered noise – oriented	of uniform mean level; mean is hidden by noise texture	2D Fourier plane. Radial width is set to one octave to match the Gabor receptive field.	phase	BW set to 30 deg to match psychophysics. Radial BW set to one octave to match psychophysics. Facilitation effect expected	with more stimulation as the coverage of the stimulus field is complete (no flat spaces)	the experiment
Bandpass filtered noise – isotropic *	No regions of uniform mean level; mean is hidden by noise texture	Ring in the Fourier domain if an isotropic filter is used;	Incoherent phase	No facilitation effect found in transducer function (i.e., only masking)	Same as above	Filtered to only contain frequencies between 0.8 and 2.0 cy/deg.
Pink Noise 5	No regions of uniform mean level; mean is hidden by noise	Ring in the Fourier domain as above if an isotropic filter is used; Rectangular window like a picture frame if Cartesian-separable filtering is used	Incoherent phase	No facilitation effect found in transducer function (i.e., only masking)	Same as above	Not used in the experiment
Selected images from a Laplacian pyramid	Mean level is arbitrary for all the bandpass images; allows for independent manipulation of mean and contrast	Like above, but rings would now by rectangular window frames in most Laplacian decompositions	Mixtures of coherent phase at step edges, and incoherent phase in complex textures	Good masking data available from Chandler, others	Same as above, but the phases represent that found in natural imagery. Since content is recognizable, there are additional cognitive effects relating to scene understanding and illumination	Images would be bandpass filtered like in the above case, but derived from natural imagery. Not used in the experiment
LPF'd natural images 6 *	Mean level determined from natural imagery's actual mean luminance but can be shifted	LPF; could be done either rotationally symmetry or Cartesian separable	Mostly incoherent phases	Good masking data available from Chandler, others	Same as above; Little facilitation; little input accommodation system	Filtered to show only frequencies between 0 and 0.6 cy/deg.
MPF'd natural images; a.k.a. bandpass filtered 7 *	Mean level is zero after filtering but can be adjusted	Rotational symmetry or Cartesian separable	Same as above		Same as above regarding phases and cognition; More signal input to the accommodation system than LPF imagery	Filtered to show only frequencies between 0.2 and 0.6 cy/deg.
HPF'd natural	Mean level is zero after	Same as above	Same as above		Same as above regarding phases	Filtered to show only

images 8 *	filtering but can be adjusted				and cognition; Has the strongest signal input to the accommodation system	frequencies between 0.8 and 2.0 cy/deg.
Natural images 9 *	Same as above	1/f^n power spectra, which is still LPF; less LPF than stimuli 6		Good masking data available from Chandler, others	Same as above regarding phases and cognition except more coherent phase due to presence of step edges; Good input to accommodation	
Signal	Luminance	Frequency characteristics	Phase characteristics	Psychophysical aspects	Neurophysiological aspects	Other comments

Obs.: * marks stimuli that were used in our study.

The flat field is strongly affected by the surround, so the field of view (FOV) of the displayed image is expected to be an important factor. A larger FOV will reduce any effects of the surround. We believe dark surround is the most stable case, since it minimizes screen reflections. The crispening effect (at boundary of image and surround) is likely the most complex phenomenon involved in a flat field. As known from Stevens work, even with this simple stimulus, the state of overall adaptation affects the nonlinearity relating perceived brightness to physical luminance. Since the FOVs are relatively large in the experiment for all cases, we expect the light adaptation level to be driven by the displayed image, as opposed to the dark ambient illumination.

The single Gabor is the next most complex stimulus as it stimulates a limited set of cortical cells, that is, those tuned for its frequency and orientation, and at that position. The stimulus spans multiple luminances both above and below the mean level. Contrast is easily controlled, and we used a Michelson based contrast (which extends to RMS contrast). This is a mean preserving contrast, as opposed to Display Contrast (Lmax /Lmin) in which the contrast adjustment also affects the luminance. It is known as the most detectable signal in terms of lowest contrast energy [1]. That is, as integrated across spatial position. Physiologically, it corresponds to the receptive fields of V1, and its position in the cortex relative to the visual field is affected by the cortical magnification factor (CMF). While the equation for a sine wave (alternative way to convey a narrow frequency band) is simpler, it is actually a more complex stimulus since it activates a wider portion of the visual field, and there are thus pooling/summation effects. Additionally, for sine waves it

is hard to estimate overall brightness without choosing either the max or min lobes as a cue.

Next is a field of Gabors with a single radial frequency and orientation, which takes the limited excitation of tuned cells and increases the excitation strength by distributing it across the cortex in accordance to the position in the visual field, as affected by the CMF. The sensitivity to this stimulus increases due to pooling (but would be less in terms of contrast energy sensitivity). In terms of the transducer function in the V1 simple cells, it is known that stimuli of this type show the effects of both masking and facilitation if the phase is held constant across the Gabors. If the Gabors are widely spaced, a subject may just ignore them and judge based on the flat regions of mean luminance between the Gabors. The spacing between them can be reduced to avoid that, but is difficult as their alignment tends to result in low contrast regions, unless a variable phase is used. If a variable phase is used the facilitation effects have not been found to occur in psychophysical studies. Further expanding the stimulated cortical area would be a field of Gabors of single base frequency but with variable orientation. This will create a ring[1] in the Fourier domain whose radius is the radial frequency. Fewer orientations would lead to a lumpy ring, and more orientations would lead to a smoother ring, such as could be achieved with random orientations. This would stimulate the entire region of V1 simple cells for that frequency, that is, all 'columns' of orientation over the CMF regions corresponding to the FOV. This approach also has the effect of reducing the interstitial regions of flat mean luminance that could be used as a cue for brightness by making the Gabors overlap.

Another approach to have similar frequency as the two Gabors cases would be to use bandpass filtered noise, where the bandpass filter is designed to simulate the cortical frequency response, whether for a single orientation, or for all orientations (i. e., isotropic). While these techniques are most commonly used, their spectra is similar to the fields of Gabors for limited orientation and the isotropic cases described above.

Low-pass filtered noise, or pink noise, would stimulate across all frequencies and orientations, but with a concentration toward the lower frequencies, in accordance with the shape of the low-pass filter. This would

[1] If radially symmetric filtering is used. If Cartesian-separable filtering is used, it would be a rectangular frame in the Fourier plane.

also be equivalent to using the n-lowest layers of a Laplacian pyramid decomposition of a noise image. The filter used is isotropic.

Another approach to generate imagery of narrow band spectral energy is to filter an optically captured image instead of noise. The filters are chosen to match the spectra of the arrays of Gabors, so most of the psychophysical and physiological aspects of V1 cortex and masking remain the same, but there is additional cognitive information because object shapes and textures are recognizable due to the phase spectra. There is less known about the cortical processing of these aspects. A similar approach could take advantage of an existing Laplacian Pyramid decomposition of an image, by taking specific layers in isolation. However, we opted to use direct bandpass filtering in order to have more precise control of the filtered frequencies. We tested four cases of this using one optically captured image: low pass, band-pass (i.e., middle range), high-pass, and unfiltered. The LPF and BP cases have mixtures of coherent and incoherent phases, with less coherent phase since steps edges are not present. In addition, the phases and frequency spectra will be varying locally throughout the image. This means facilitation effects should be minor. Also, due to their lack of high frequency energy, there will be a weak accommodation response to these stimuli. This means the eye will be less likely to be exactly focused on the screen. The edge of the screen will still act as an accommodation input, of course, but for the wider FOVs studied, it will be a weaker input signal unless the viewer specifically foveates the screen edge. The HPF image will contain mixtures of coherent and incoherent phase, also varying locally throughout the image and frequencies. It will have the strongest accommodation input of all the tested images. The unfiltered image would have a spectrum similar to the 1/f (pink noise) stimuli used partially due to the optics of capture, and partially due to natural image and civilized image statistics [2]. However, there are much more complex phase relationships. It will have a sufficient accommodation input, second only to that of the HPF image.

Perceptual study:

Our study was conducted using the following stimuli, marked with a red asterisk * in the table above. Cutoff frequencies for Fourier domain filtering are described in the table where appropriate. Bandpass filtering was achieved using a Butterworth filter.

• 1 Flat field;

• 2 Single Gabor pattern;
• 3 Field of smaller Gabors with the same orientation (45◦) and random phase;
• 4 Field of smaller Gabors with random orientations;
• 5 Pink noise (1/f);
• 6 Low-pass of a natural image;
• 7 Bandpass favoring low frequencies of the same image;
• 8 Bandpass favoring the highest response of the CSF of the same image
• 9 Unprocessed grayscale natural image.

The natural image was carefully selected to fit specific criteria:
• It must not contain large empty areas (to avoid participants making decisions on "flat" background regions),
• Needs to be nearly symmetric in the luminance domain around the mean in order to appear being at full contrast after going through our display pipeline (this implies it shouldn't contain large dark patches in any color channel),
• Should contain a large number of textured regions,
• The image must have consistent lighting since we did not want to invoke effects caused by inconsistent lighting [3] as can occur in image compositing.

For the main part of the experiment, the stimuli had their average luminance set to 10 nit, with amplitude 10, resulting in images ranging from 0-20 nit. The resolution was always set to 1920x1080 pixels. The stimuli were generated with the desired per pixel nit values first, followed by the application of the appropriate gamma function and corrections for the target display.

Citations:

[1] Watson, Andrew B., H. B. Barlow, and John G. Robson. "What does the eye see best?." *Nature* 302.5907 (1983): 419-422.
[2] Graham, Daniel J., and David J. Field. "Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities." *Spatial vision* 21.1 (2007): 149-164.
[3] Tan, Minghui, et al. "The perception of lighting inconsistencies in composite outdoor scenes." *ACM Transactions on Applied Perception (TAP)* 12.4 (2015): 18.