Title:

Immersion in Modified Reality: Sensitivity to the Slope (α) of the Amplitude Spectrum is Dependent on the α of Recently Viewed Environments

Short Title: Sensitivity to α in Modified Reality

Authors:

Bruno Richard[1]

Patrick Shafto[1]

Author Affiliation:

[1]Department of Mathematics and Computer Science, Rutgers University - Newark, Newark, New Jersey, USA, 07102

Corresponding Author: Bruno Richard, Department of Mathematics and Computer Science, Rutgers University - Newark, 101 Warren Street, Rm 216, Newark, New Jersey, USA, 07102, Email: bruno.richard@rutgers.edu

Keywords: Natural Scenes, Amplitude Spectrum Slope, Head-Mounted Display, Modified Environment, Bayesian Observer Model, Multimodal Prior

Abstract

Scenes contain many statistical regularities that, if accounted for by the visual system, could greatly benefit visual processing. One such statistic to consider is the orientation-averaged slope (α) of the amplitude spectrum of natural scenes. Human observers are differently sensitive to αs, and they may utilize this statistic when processing natural scenes. Here, we explore whether discrimination sensitivity to α is associated with the recently viewed environment. Observers were immersed, using a Head-Mounted Display, in an environment that was either unaltered or had its average α steepened or shallowed. Discrimination thresholds were affected by the average shift in α: a steeper environment decreased thresholds for very steep reference αs while a shallower environment decreased thresholds for shallow values. We modelled these data with a Bayesian observer model and explored how different prior shapes may influence the ability of the model to fit observer thresholds. We explore three potential prior shapes: unimodal, bimodal and trimodal modified-PERT distributions and found the bimodal prior to best-capture observer thresholds for all experimental conditions. Notably, the prior modes' position was shifted following adaptation, which suggests that a priori expectations for α are sufficiently malleable to account for changes in the average α of the recently viewed scenes.

Author Summary

Human observers are sensitive to the statistics of natural scenes. Here, we investigated whether differences in discrimination thresholds to the amplitude spectrum slope (α) of natural scenes is tied to the distribution of α in typical visual environments of humans. We measured discrimination sensitivity to α within the environment of observers in a Head-Mounted Display, which overlaid our stimuli onto a real-time recording of the observer's environment. Observers were then immersed in a modified environment; the α of each image was steepened or shallowed to shift the distribution of α observers were exposed to. Our adaptation procedure altered discrimination thresholds as they showed an additional peak in sensitivity for α values near the new average α of the modified environment. Our Bayesian modelling indicates that the prior for α is bimodal and sufficiently flexible to adapt to novel environments with different α distributions. Overall, our findings demonstrate that the human visual system continuously adapts to the current statistical structure of its natural environment.

Our visual world is diverse but contains spatial regularities that span across the various environments we encounter. There is evidence that accounting for these spatial regularities in encoding can improve the visual system's efficiency in processing complex scene properties [1–4]. One such regularity is the slope of the amplitude spectrum (α) of natural scenes, which defines the association of luminance contrast to spatial frequency (f) $1/f^{\alpha}$. On average, the α value of natural scenes is 1.08 [5–11]. Observers are sensitive to and influenced by α: they exhibit different discrimination and recognition abilities for different α values [1, 12–14] and can adapt to α [15–17]. Biases in the discrimination ability of observers for α may be associated with the recent visual experience of observers [15]. Here, we explore how the discrimination sensitivity for α is related to prior expectations of the distribution of αs gathered from recent perceptual experience. We used a Head-Mounted Display (HMD) system paired with a camera to display and modify the natural visual environment of observers and measure discrimination thresholds to α. Finally, we interpret our findings according to a Bayesian observer model [18–20] and derive the shape and malleability of the prior for the αs of natural scenes.

Measuring the discriminability of α has been a popular method for determining the shape of the internal representation for α [9, 21–23]. For example, a common finding when measured with noise stimuli, modified to have a particular α value, is that the discriminability of α is best (thresholds are lowest) when the reference α is between 1.0 and 1.3, and rises as the α is steepened or shallowed [13, 24, 25]. Peak discriminability matches that of the average α of natural scenes. Thus it appears that the human visual system is best able to discriminate the spatial characteristics of scenes it encounters often. However, peak in discriminability can depend on the original α value of the image. When observers are asked to discriminate from the original α of a natural image, thresholds are significantly higher than when observers complete the discrimination task from an image with a modified α[1] [26]. Tolerance for small deviations in the natural α of an image [22, 23, 26] may assist in optimizing visual processing by facilitating the discrimination of various objects in the natural world [12, 14]. Thus, while the human visual

[1]Note that modifying the α of a natural image will alter its perceived focus: steepened images will appear blurred while shallowed images appear sharpened.

system peak sensitivity to α is for values regularly encountered ($\alpha \approx 1.0$), it is simultaneously tolerant for minor deviations from the natural α of an image, which suggests some process of adaptation to the spatial characteristics of the environment.

Visual adaptation is the process by which exposure to a stimulus affects visual processing for subsequent stimuli [27–30]. When the adapting stimuli are spatially complex — noise stimuli modified to have a natural α value — adaptation acts to renormalize the percept of observers towards an internal norm; a neutral point in the responses of the detecting mechanisms [15, 28]. Adaptation to artificially steepened or shallowed images generated a shift in the perceived α of a new image in the opposite direction to that of the adaptor α, No shift in the perceived α of images was measured when the adapter had an α of 1.0, which is coincidentally the average value of α across all natural scenes [15]. It appears that through visual experience, the human visual system has developed an internal representation or an expectation for α values commonly encountered.

The statistical structure of natural scenes is known to influence biases in sensitivity to visual features [19, 20, 31, 32]. For example, there is an orientation anisotropy in human vision whereby sensitivity to horizontal content embedded in natural scenes is worse than that of obliques (i.e., the horizontal effect; [33–35]). This bias in orientation sensitivity is believed to stem from the over-representation of horizontal content in natural scenes [11, 19, 33], which is perceptually normalized by reducing sensitivity to horizontal content. Indeed, if observers are immersed in an environment that does not contain the natural over-representation of horizontal content, the strength of the horizontal effect diminishes severely [31]. These effects occur relatively rapidly (within an hour or two of immersion), which is a strong indication that the expected structure of natural scenes can be relearned [31, 32]. The change in sensitivity to orientation contrast was well-explained by a change in the prior of a Bayesian observer model which matched the new orientation distribution of the modified environment [31]. If the internal representation for α is determined by the distribution of commonly viewed αs in the natural environment, then it may be able to adjust given a novel environment statistical structure, as has been demonstrated with orientation sensitivity in natural scenes.

Here, we explore whether sensitivity to α is associated with the statistical structure, or distribution, of α in natural scenes by measuring α discrimination thresholds prior to and following immersion in an environment that no longer contains the typical α distribution of scenes. Observers were immersed in the environment with a Head-Mounted Displays (HMDs) and a camera that recorded and displayed the visual world in near real-time (a display system that we have named Modified Reality; [31]). All experimental conditions, the measurement of α discrimination thresholds and the adaptation to the modified environment were conducted in the Head-Mounted Display. This method greatly facilitates the measurement of sensitivity to features in the natural environment of observers while maintaining sufficient experimental control on the presented stimuli. The environment was modified by steepening or shallowing the distribution of αs. Finally we develop a Bayesian observer model to explore the possible shape of a prior for α and verify if the prior can adjusts to match the novel environmental distribution of α [18, 19, 31].

Results

Average α discrimination thresholds measured at baseline and following adaptation conditions (+0.4 and -0.4) are shown in **Figure 1**. α discrimination thresholds measured in the Oculus Rift DK2 follow the typical peak and trough commonly observed with more traditional measurement tools (i.e., CRTs) [13, 21, 22, 24, 25].

The magnitude of change between the baseline and adaptation conditions was quantified with the effect size Hedge's g. We observed little change in discrimination thresholds to most reference αs used in this study ($g_{\alpha_{0.4}} = -0.141$, $g_{\alpha_{0.7}} = -0.273$, $g_{\alpha_{1.0}} = -0.162$,$g_{\alpha_{1.6}} = 0.189$), but did measure larger facilitation effects to steep reference α of 1.3, $g_{\alpha_{1.3}} = -0.672$, and 1.9, $g_{\alpha_{1.9}} = -0.482$. Adaptation to a shallowed environment (-0.4) generated a much more compelling change in tuning to α. Adaptation selectively reduced thresholds for a reference α of 0.7, $g_{\alpha_{0.7}} = -1.518$, and 1.3, $g_{\alpha_{1.3}} = -1.128$, while leaving thresholds for all other αs unaffected ($g_{\alpha_{0.4}} = -0.350$, $g_{\alpha_{1.0}} = -0.447$, $g_{\alpha_{1.6}} = -0.264$, $g_{\alpha_{1.9}} = -0.376$). The behavioral change following exposure to modified environments suggests some influence of the environmental αs on

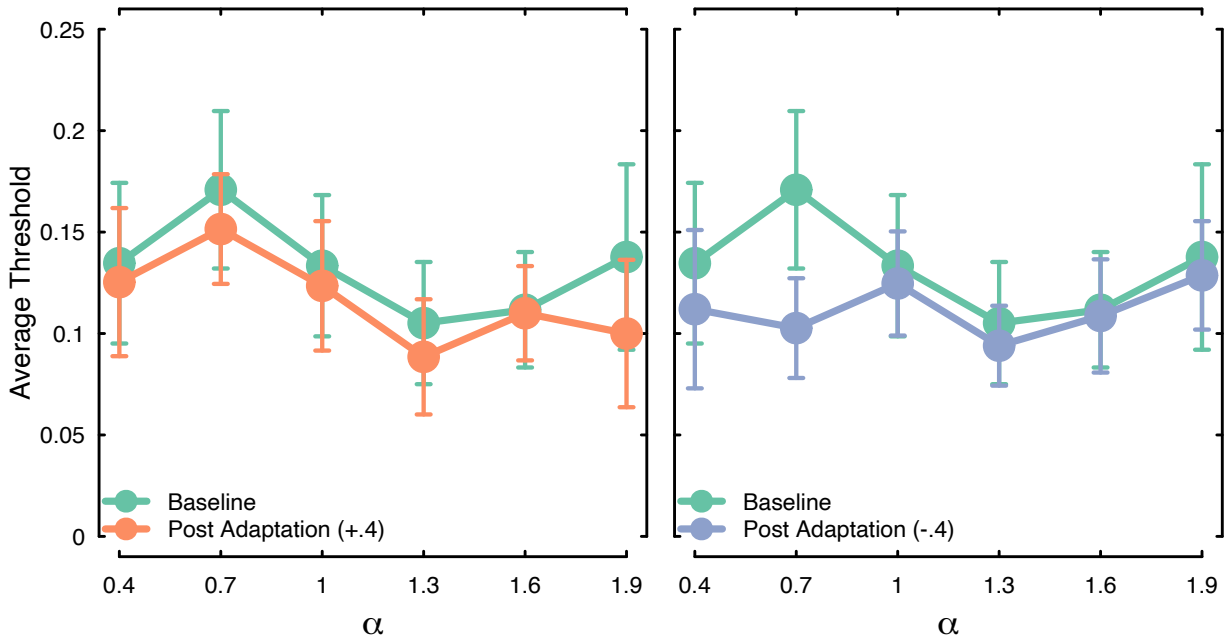

Figure 1: Average discrimination thresholds at the five reference αs measured in this study across all three observers for baseline (green), post steep adaptation (orange) and post shallow adaption (blue). Error bars represent $\pm$ 1 standard error of the mean.

human sensitivity to α. When observers are immersed in a world that is shallowed, where the average α of the environment is lowered, discrimination sensitivity to shallow α that human observers are not normally sensitive to increases by 1.5 standard deviations. A similar effect is found for the steepened environment, albeit to a lesser extent.

To explore this further, we built an observer model that uses a Bayesian estimator to capture α discrimination thresholds prior to and following adaptation. If the change in discrimination we observed is related to a change in the internal representation for α, then these effects should be well-described by a change in the prior probability density function of a Bayesian observer. Indeed, previous work has demonstrated the success of this approach in defining the prior for orientation contrast in natural scenes [19], and changes in the observer prior following exposure to environments with no natural bias in orientation contrast [31]. We built our prior distribution according to empirically derived environmental αs distributions [5–7, 9, 11]. An interesting factor to consider is that the mode of α distributions varies according to the type of environment they represent (e.g., human-made versus natural) [6]. Whether the different modes of α distributions

are represented in the human visual system is unclear. For example, a prior for α could be defined as a single unimodal distribution centered on $\alpha = 1.0$, a commonly measured distribution when images from all environments are combined [5, 6, 8]. The prior could also be multimodal in a effort to better represent the various environments humans observers can encounter. We build three different priors that differed in their number of modes (unimodal, bimodal, and trimodal) and verify which best performed in predicting observer discrimination thresholds. The α priors, $p(\alpha)$, were defined as a modified-PERT distributions. The PERT distribution is a modified Beta distribution, defined as

$$p(\alpha) = \frac{(\alpha - a)^{\delta-1}(c - \alpha)^{\epsilon-1}}{Beta(\delta, \epsilon)(c - a)^{\delta+\epsilon-1}} \tag{1}$$

with parameters a, b, and c that represent the minimum, mode and maximum values of the distribution, respectively. The two parameters of the Beta distribution are

$$\delta = 1 + \gamma \frac{b - a}{c - a}, \tag{2}$$

$$\epsilon = 1 + \gamma \frac{c - b}{c - a} \tag{3}$$

and include an additional parameter (γ) which controls how much weight (probability) is assigned to the mode of the distribution and is important for skewed distributions where the distance between the lower limit (a) and the mode (b) or the upper limit (c) and the mode (b) are different.

The observer likelihood and posteriors were built in an identical fashion to the methods described in Girshick et al., (2011) and Schweinhart et al., (2017). We constructed a 2D matrix where the horizontal axis represented the stimulus α and the vertical axis, the measured α . The measurement distribution $p(m|\alpha_i)$ was defined as a Gaussian centered on the stimulus α,

$$p(m|\alpha) = \frac{1}{\sigma_\alpha \sqrt{2\pi}} e^{-\frac{(\alpha - \bar{\alpha})^2}{2\sigma_\alpha^2}} \tag{4}$$

with a variance term, σ_α^2, equal to the variance of the last five reversals for each staircase collected

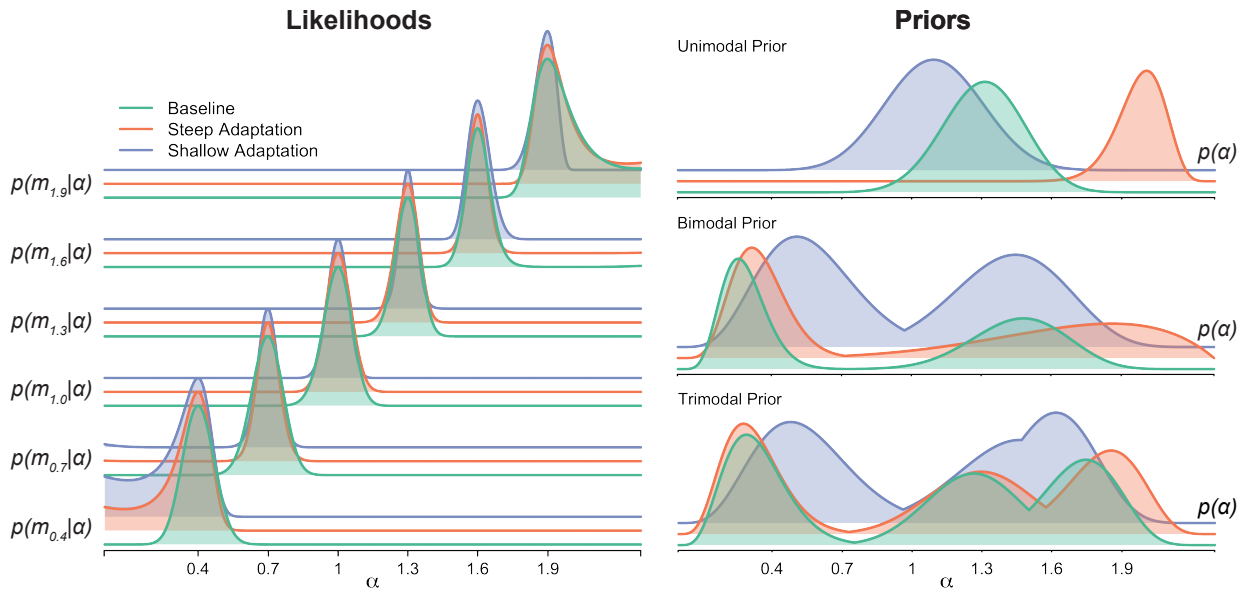

Figure 2: The resulting likelihood and prior distribution before and following adaptation. The likelihood distributions for each of the 5 reference α values used in this study. The best fitting unimodal, bimodal and trimodal prior distributions. All show a similar peak in α probability prior to and following adaptation for steeper α (> 1.0), while the bimodal and trimodal distributions add a peak at a shallower α. The trimodal prior only has a discernible third peak in the steepened condition while its shape for the shallow condition is nearly identical to that of the bimodal prior.

for each of the experimental conditions in this study [18, 19]. The posterior distribution, $p(\alpha|m)$, was computed using Bayes' rule, $p(\alpha|m) \propto p(m|\alpha)p(\alpha)$. We defined model thresholds as the distance between the estimator, built from the mode of the posterior distribution for each measurement, and a straight line along the diagonal of the matrix. Model parameters were optimized with *fminsearch* in MATLAB, minimizing the sums of square error between the model output and observer α discrimination thresholds, prior to and following adaptation (see **Figure 2** for best fitting priors).

The resulting fits of our probabilistic model to the baseline and adaptation data are shown in **Figure 3**. The result of each model prior is represented in separate subplots, fits for the steep environment are shown in the upper row while the fits to the shallowed environment are in the bottom row. The bimodal ($R^2_{baseline} = .970$, $R^2_{steep} = .829$, $R^2_{shallow} = .859$) and trimodal ($R^2_{baseline} = .982$, $R^2_{steep} = .968$, $R^2_{shallow} = .767$) priors performed well and captured most of the observer data, while the unimodal prior ($R^2_{baseline} = .292$, $R^2_{steep} = .640$, $R^2_{shallow} = .281$) was unable to fit the peak and trough characteristic of α discrimination. The bimodal and trimodal

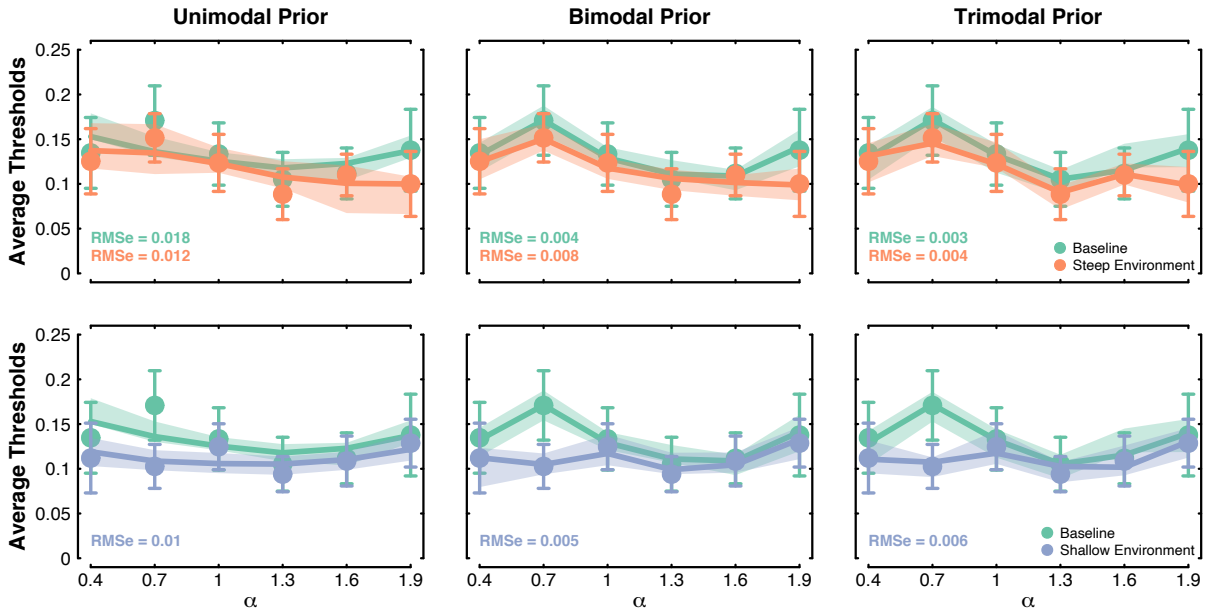

Figure 3: Resulting model fits for the unimodal, bimodal and trimodal prior probability distributions. Points are human discrimination thresholds. Lines are model predictions. Filled area are 95% confidence intervals built by sampling observer data (with replacement) and refitting the model 1000 times.

prior appear at par in their explanatory power of human α discrimination thresholds. As they differ in their number of free parameters — 4 and 6 respectively — we further compared the bimodal and trimodal prior distribution model performance with AIC scores [36] and an extra Sums of Squares test. When factoring in the number of free parameters, the bimodal prior ($AIC_{bimodal} = -17.45$) outperformed both the unimodal ($AIC_{unimodal} = -16.35$) and trimodal ($AIC_{trimodal} = -15.37$) priors. The bimodal prior also explained a statistically significantly larger portion of the variance in our data than the trimodal prior, $F(1, 13) = 11.60, p = .001$.

All three prior variants of our model exhibited similar shifts as a consequence of our adaptation paradigm (see **Figure 3**). At baseline, the prior distribution of all models had a mode near an α of 1.3 (unimodal = 1.31, bimodal = 1.48, trimodal = 1.26), with the bimodal and trimodal priors showing a secondary peak at shallow αs (< 0.4) and a third mode at a steep α of 1.74 for the trimodal prior. Following adaptation to a steepened environment, the unimodal and bimodal prior probability distribution models exhibited a rightward shift in their modes. The unimodal distribution now peaked at an α value of 2 while the bimodal prior had an upper most peak at 1.86 with little change at the second mode for shallow α. Only a slight increase at the

upper most mode of the trimodal prior was observed following adaptation to a steepened environment. A similar shift was observed after adaptation to the shallowed environment. The unimodal prior probability distribution was leftward shifted compared to baseline (mode α = 1.09). The shift for the bimodal and trimodal priors occurred in their lower peak, which was rightward shifted compared to baseline (α = 0.51 and α = 0.48, respectively).

A Bimodal Prior for α

A bimodal prior for α with a peak in probability at α values seldom encountered in the environment was unexpected. As it is, however, the best fitting prior in our model, it is worthy to explore whether this particular prior shape is indeed capable of accurately encoding the α of natural scenes and outperform the other priors we investigated above. We simulated an encoder population whose responses were taken from Gaussian distributions with centers (their preferred α) determined by the different priors used above (see **Figure 4**) [19, 37]. The encoder population was decoded according to a standard population vector decoder ([37]), which computes the sum of the response vectors for a given stimulus ($r_n(y)$) weighted by their preferred α:

$$\hat{\alpha}(y) = \frac{\sum_{n=1}^{N} \alpha_n r_n(y)}{\sum_{n=1}^{N} r_n(y)}. \tag{5}$$

We simulated the responses of the decoder to 100 external stimuli — α values from natural and human-made environments taken from distributions of α measured by Hansen and Essock (2005). We used 1000 Gaussian distributions with centres taken from the unimodal, bimodal, and trimodal priors in addition to a Gaussian prior centred on an α = 1.0, which served as an additional control condition.

The decoder response accuracy is shown in **Figure 4**. The figure displays the values of the external stimulus α on the x-axis and the decoded values on the y-axis. Circle markers represent α values for man-made environments, while the square markers represent αs for natural scenes. The decoder performed best when the bimodal prior defined the centres of our Gaussian channels as the data points lie nearly perfectly on the diagonal. The decoder combined with the trimodal

prior performed similarly well for steeper α values but overestimated shallow α values as its decoded response was above the unity line. Results of the unimodal prior and Gaussian were poor. Both overestimated α values for αs below their respective peaks and underestimated for values above their peak. A population of Gaussian channels with peaks in tuning at shallow and steep α values is best accurately decoding the α value of both natural and human-made environments, while more uniform (unimodal / Gaussian) or heterogeneous (trimodal prior) tuning distributions do not perform nearly as well. We also looked at the difference between the actual input α and that outputted by the population decoder and found it to largely resemble the α discrimination curve of human observers. It was most accurate for α values near 1.0-1.2 and worse for shallower and steeper αs.

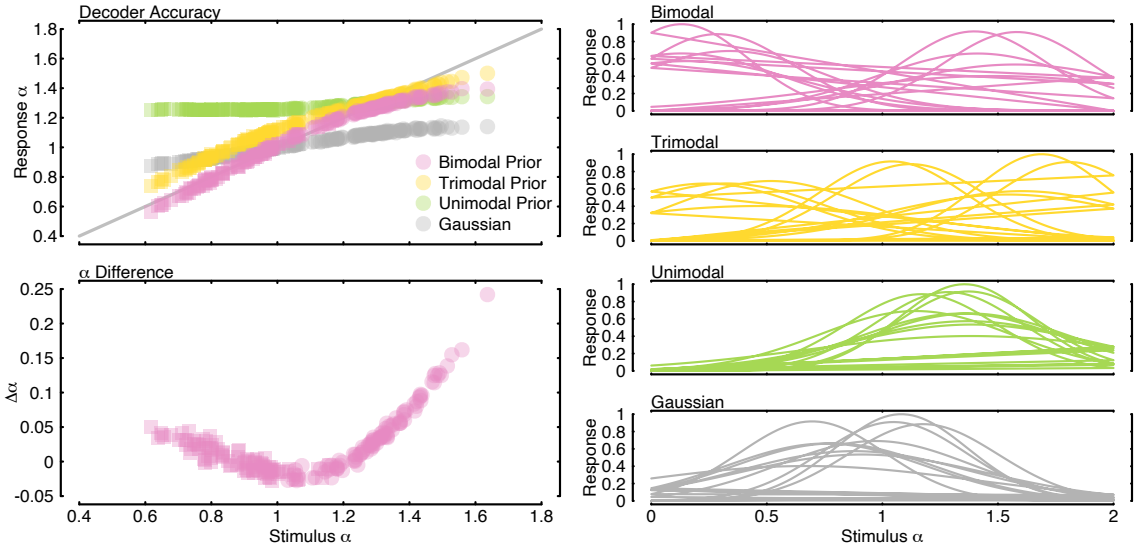

Figure 4: Results of the population encoder analysis. Decoder accuracy, shown in the top left, plots the actual α value fed into the population and the decoded α outputted by the population for all four priors. Circle markers represent α values taken from the human-made α distribution while square markers are α values taken from the natural scenes distribution. The diagonal is the unity line. Examples of the population channels built from the different priors are shown in the left column. Peaks in each distribution represent the preferred α value of each channel and are taken from the prior distributions. Finally, α difference plots the difference between the pink markers in the decoder accuracy plot and the unity line. It demonstrates how a bimodal prior generates the smallest error in decoding α for values between 1.0 and 1.2, a similar pattern to that of human observers when measuring α discrimination thresholds.

Discussion

We measured discrimination thresholds to the slope of the amplitude spectrum of noise patches before and following immersion in a modified environment, all within an HMD. We found baseline α discrimination thresholds to have the same pattern across reference αs as is typically measured in more traditional laboratory equipment. Immersion in a modified environment led to a change in α discrimination thresholds. When the average α in the environment is shallower than the original (i.e., $\alpha - 0.4$), discrimination thresholds for shallow reference α (0.7) were reduced, while thresholds for steeper α were unaffected. When the αs were steepened, we observed a slight decrease in thresholds at very steep reference α (1.9). Computationally, we explained our psychophysical results with a Bayesian observer model [18] similar to previous studies investigating the statistics of natural scenes [19, 31]. We found the best fitting model to have a bimodal prior for α and modes at αs of 0.25 and 1.48. These values approximate but do not match peaks in the natural distribution of α. We verified if this prior could have a potential functional relevance by building an encoder population with tuning peaks taken from our investigated priors. The encoding population was taken from the bimodal prior performed best and generated decoding patterns that mirror human sensitivity to α. Finally, as has was found with other visual features [31, 32, 38, 39], the modes of the bimodal prior are not fixed and adapt to the environment of the observers. Specifically, the upper mode of the prior was rightward shifted following immersion in the steepened world. In contrast, the lower mode of the prior shifted to steeper α values after immersion in the shallowed world. Changes in the peak probabilities of the prior for α were sufficient to capture the changes in sensitivity to α following immersion in the modified environment.

We explored three possible prior probability density functions in our modelling approach: a unimodal, bimodal, or trimodal modified-PERT distribution and found the bimodal prior to fit the observer thresholds best overall. A bimodal prior is not an exact match to the environmental distribution of α, which is typically unimodal with a different peak position based on the

environment type (natural VS human-made). At baseline, the peak probabilities of our bimodal prior were a very shallow α of 0.25 and a steep α of 1.48, slightly steeper than the mode of the α distribution of the testing environment (α = 1.29; see **Figure 5**). Thus, we appear to find a prior shape primarily motivated by functional accuracy instead of being an exact match to the environmental α. The population encoder analysis [37] demonstrated that a slightly heterogeneous set of Gaussian channels for α, with peaks at shallow and steep α values, generated the most accurate decoding results. The decoding accuracy or the difference between the external α fed to the population and the decoding output closely resembled the human discrimination threshold tuning curve. Increasing the heterogeneity of the encoding channels by constructing them according to probabilities from the trimodal prior hampered decoding accuracy. When the Gaussian channels were built from priors that match the environmental distribution of α, (the unimodal and Gaussian priors), leading to a homogeneous set of channels, decoding accuracy was poor. Thus, a prior that matches the environmental distribution of α is incapable of accurately capturing the behavioral performance of our observers on the discrimination task or accurately encode the α of scenes. This is unlike the prior for orientation contrast, which does match the biases in contrast energy across orientation [19, 40]. While a prior that matches the environmental condition is believed to be more optimal in the encoding of stimuli [31, 40], our results suggest that the shape of the prior may also stem from functional necessity.

Our prior for α may not perfectly represent the natural distribution of α in the environment, but this does not mean it cannot be influenced by the current statistics of the visual world [31, 32, 38, 39]. Indeed, we found that a change in the statistics of scenes (a shift in α) generated a shift in the position of both peaks in our bimodal prior. A steep environment, which had all of its α values increased by 0.4, generated a rightward shift of the steep mode while leaving the peak of the shallow mode unaffected. The shallow environment, which had all of its α decreased by 0.4, led to a rightward shift of the shallow mode of the prior probability distribution and did not affect the steep mode. However, it is interesting to find that immersion in a shallowed environment shifted the lower peak of the bimodal prior upwards and nearer to the average α of the modified world (α

= 0.89). This was not the direction of the shift we had anticipated, as it does not necessarily reflect the change in the environment. However, it is the prior that best fits observer discrimination thresholds. The decrease in discrimination thresholds to a reference α of 0.7 appears tied to the rightward shift in the lower peak of the prior. Indeed, changes in the prior alone were sufficient to explain the change in observer α discrimination thresholds recorded before and following immersion in the modified environment. We found little to no change in the observer likelihood (see **Figure 2**). In this way, our findings differ from other studies that have reported adaptation effects as being attributed to changes in the likelihood of observers, such as the tilt aftereffect [41]. The difference in our findings can likely be attributed to the duration of adaptation. Our modified reality procedure subjects observers to the adaptation environment for an extended period (1 hour + 45 minutes of subsequent testing), which is significantly longer than traditional adaptation studies that adapt for periods of seconds or minutes. While short adaptation periods are unlikely to generate any meaningful change in the prior of observers, prolonged immersion in modified worlds [31] or exposure to stimuli [39] does stimulate the learning of novel stimulus properties / environmental statistics. Learning can generate a systematic change in the prior of observers that mirrors the change in the environment [32], or as we observed for the α of natural scenes, generate a change that appears to be more functional in nature.

Finally, it is worthy to note that psychophysical measurements of discrimination thresholds for the slope of the amplitude spectrum are traditionally completed in laboratory environments with monitors (e.g., CRTs) that are large and of a sufficiently high resolution to display a wide range of spatial frequencies [9, 13, 24, 25]. Traditional CRTs appear not to be necessary when measuring sensitivity to the slope of the amplitude spectrum of natural scenes. While implementing psychophysical paradigms in HMDs is not novel [31, 42], the measurement of α discrimination threshold, which is dependent on the spatial frequency resolution of the display, is novel. We have demonstrated that discrimination thresholds to α can be measured accurately with the smaller screens of HMDs limited in the range of spatial frequencies they can display. The spatial frequency resolution of the Oculus Rift DK2 only reaches 5.5 cycles/^{o} of visual angle.

Still, α discrimination thresholds measured in the Oculus Rift DK2 showed the expected peak and trough for reference αs of 0.7 and 1-1.3, respectively. The peak and trough were also of similar magnitude to thresholds measured in other studies (using the same staircasing methodology) [24, 25]. Human observers do not need a wide range of spatial frequencies to discriminate α and complete the task with relatively limited spatial frequency content.

Conclusion

Sensitivity to α was well-explained using a Bayesian framework with a bimodal prior with peaks in probabilities at very shallow and steep αs. This particular prior shape is not fixed, which brings additional support to the finding that the representation of the distribution of statistical properties of natural scenes is a dynamic process that continuously adjusts to the current environment of the observer [31, 32, 38, 39]. It, additionally, appears to increase discrimination sensitivity to the new common αs of the environment. A bimodal prior for α is not an exact representation of its natural distribution but is a property that allowed a population encoder-decoder analysis to generate an output that was very similar to that of our human observers. Specifically, we found that a non-uniform set of encoding channels that prefer steep and shallow αs, instead of being centered on α values commonly encountered in the environment, was best for decoding α values taken from natural and human-made images. This is indicative of an optimal encoding of the distribution of contrast energy across spatial frequency by the human visual system [19, 32, 43], and suggests a functional relevance of α in the processing of the observer's environment.

Materials and Methods

Participants

Three male observers (age 29, 33, and 34) took part in this study. All participants were experienced psychophysical observers with prior training on the amplitude spectrum slope discrimination task and had normal or corrected to normal visual acuity. Informed consent was obtained from all participants prior to them taking part in the study.

Apparatus

The experiment was conducted in an *Oculus Rift Development Kit 2* (DK2). The DK2 has two 5.7 inch OLED displays, each having a monocular resolution of 960x1080 pixels and a refresh rate of 75Hz. The Field of View of the Oculus Rift DK2 is 100°, resulting in a visual resolution of approximately 11 pixels per degree of visual angle [2]. The observer's environment was captured by a *See3Cam_130* USB 3.0 camera mounted on the Oculus Rift. The camera continuously collected images of the observers' environment at a rate of 60 fps during the experiment. The images recorded by the camera were sent to the laptop computer (Dell Latitude 3470), using the MATLAB Image Acquisition Toolbox, filtered in MATLAB and displayed back to the Oculus Rift with the the Oculus VR library of *Psychtoolbox* [44, 45].

Stimuli

The environment images collected by the mounted camera were converted to grayscale and cropped to a size of 720x1080 pixels. The 120 pixels of the Oculus display above and below the presented environment were set to mean luminance (RGB [128,128,128]). At baseline condition, only the RMS contrast of the environment images was adjusted to match that of the target stimuli (RMS = 0.15). For the adaptation condition, images collected from the environment had their global α changed by either +0.4 or -0.4 (see **Figure 5**). The environment α was changed in an

[2] This low visual resolution is a consequence of having displays with a large field of views very close to the observers' eyes, $\sim$ 5cm.

identical manner to the creation of the synthetic noise images used to measure discrimination thresholds (**Figure 5A**). The amplitude of each spatial frequency coefficient was divided by the average amplitude (calculated across orientation for that spatial frequency) to create a flat ($\alpha = 0.0$) amplitude spectrum [22]. In this form, the α of the environment image was adjusted by multiplying each spatial frequency's coefficient by $f^{-\alpha_{adjust}}$ (the original image α + 0.4 for the steeper condition or - 0.4 for the shallower condition). The RMS contrast of α_{adjust} images was then adjusted to 0.15.

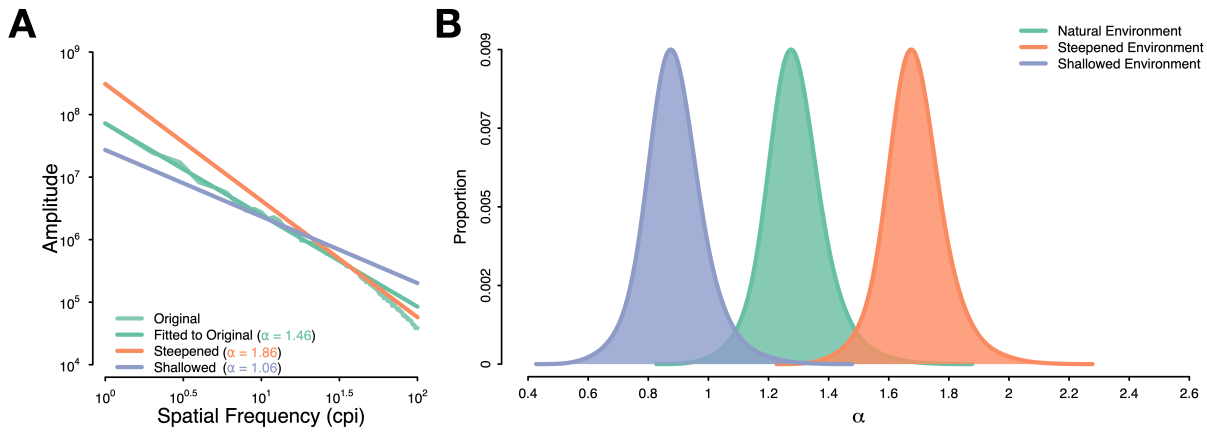

Figure 5: **A** α modification procedures: We fit a linear regression to the log orientation averaged amplitude spectrum of the original image (Original) to measure its $\alpha_{original}$. The new amplitude spectrum slope was created by first dividing the spatial frequency coefficients by the average amplitude to create a flat amplitude spectrum. The α of the original image could then be adjusted by multiplying each spatial frequency coefficient by $\alpha_{original} + 0.4$. The modified spectrum was then scaled vertically to ensure that the total contrast energy in the modified image was identical to that of the original. **B** The distribution of α values in the environment of our observers. The α distribution of our laboratory was similar to that reported by [6] for fully carpented environments with a mean α = 1.29 and standard deviation of 0.09. The steep environment distribution has a mean of 1.69 while the shallow environment had a mean of 0.89.

All test stimuli consisted of synthetic visual noise patterns constructed in the Fourier domain using MATLAB (Mathworks, Natick, MA) and corresponding Signal Processing and Image Processing toolboxes. We opted for noise patterns with different α values as test stimuli, and not natural images, as these were overlaid onto the modified environment when we measured discrimination thresholds and this greatly simplifies the identification of the test stimuli for observers. The visual noise stimuli were created by constructing a polar matrix for the amplitude spectrum and assigning all coordinates the same arbitrary amplitude coefficient (except at the

location of the DC component, which was assigned a value of 0). The result is a flat isotropic broadband spectrum (i.e., α = 0.0), referred to as the template amplitude spectrum [22, 24]. In this form, the α of the template spectrum can be adjusted by multiplying each spatial frequency's amplitude coefficient by f $^{-\alpha}$. The phase spectra of test stimuli were constructed by assigning random values from $-\pi$ to π to the different coordinates of a polar matrix while maintaining an odd-symmetric phase relationship to maintain conjugate symmetry. The noise patterns were rendered into the spatial domain by taking the inverse Fourier transform of an α-altered template amplitude spectrum and a given random phase spectrum. The phase spectrum for all stimuli presented within a trial was identical but randomized from trial-to-trial. The $1/f^{\alpha}$ stimuli were 20° in size (220 pixels). RMS contrast (the standard deviation of all pixel luminance values divided by the mean of all pixel luminance values) was fixed to 0.15.

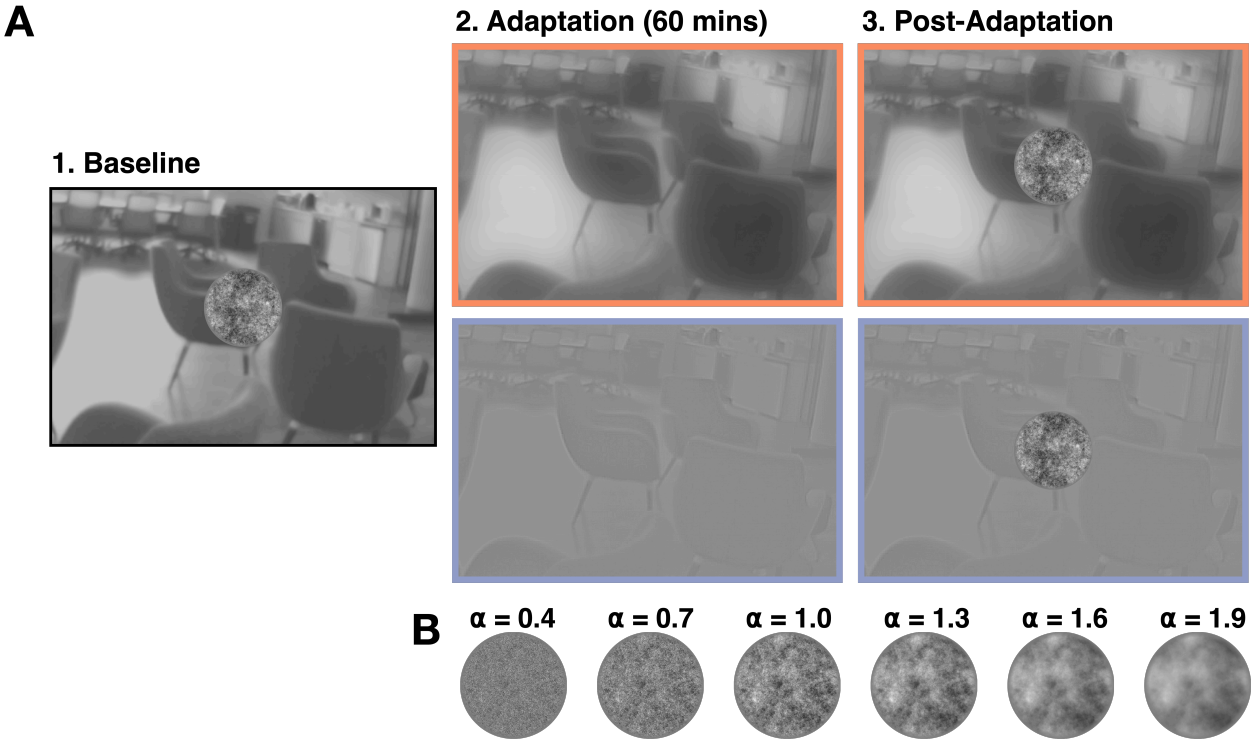

Figure 6: **A** Prior to adaptation, participants completed a baseline α discrimination task in their unaltered (gray-scaled and RMS contrast adjusted to 0.15). Participants did 2 staircases per reference α [0.4, 0.7, 1.0, 1.3, 1.6] (see **B** for example reference α). Participants then began the adaptation procedure, where they viewed the modified world (global α of the environment were steepened by +0.4) for a period of 10 minutes. Following adaptation, participants completed the same α discrimination task as baseline, but with the modified environment in the background. Note that the contrast of the test pattern has been increased for visibility.

Procedures

The experimental procedures measured sensitivity to α prior to and following adaptation to a modified environment. Participants first completed an α discrimination task with $1/f^{\alpha}$ noise stimuli placed atop the unaltered environment (RMS adjusted only; see **Figure 6**). Discrimination thresholds were estimated by a temporal three-interval, two-alternative "Odd-Man-Out" forced choice task [13, 24, 25]. The trial-to-trial change in the image's α was controlled by a 1-up, 2-down staircase procedure targeting 70.71% correct [46, 47]. The staircase began with a difference between the odd and reference stimulus of 0.5, which decreased in linear steps (step size = 0.02) towards the reference α when the observer made two consecutive correct responses. The difference was increased by 0.02 when observers made an incorrect response (1-Up / 2-Down Rule).

Each trial began with a black (RGB [0, 0, 0]) fixation point (0.5^{o}) presented for 1000ms at the center of the display, followed by three stimulus presentation intervals (250ms each) interlaced by a blank screen (environment only) for 500ms. The second interval always contained the reference α while either the first or third contained the odd stimulus (the other being identical to the reference α interval). Participants indicated which interval, either the first or the third, they perceived as being the "Odd-Man-Out" via keyboard press. The staircase continued until 12 reversals occurred and thresholds were estimated by averaging the α values of the odd stimulus for the last 5 reversals. Observers completed two staircases per reference α value [0.4, 0.7, 1.0, 1.3, 1.6, 1.9]; the discrimination thresholds of each were averaged for data analysis.

Immediately following baseline, participants completed a 60 minute adaptation period immersed in the modified environment (either α shifted by +0.4 or -0.4) in the HMD. Participants were instructed to look around their environment and interact with objects and individuals present around them but to remain seated or standing in place. Observers then repeated the α discrimination task with the modified environment presented in the background. Participants repeated the experiment with the other modified environment (either α shifted by +0.4 or -0.4) no less than 24 hours following the first experimental session. Each experimental session took

approximately 2.5 hours to complete (45 minutes per α discrimination task and 60 minutes of adaptation).

Acknowledgements

We would like to thank Bruce C. Hansen for his insightful comments on this manuscript.

References

[1] David J. Field and Nuala Brady. "Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes". In: *Vision Research* 37.23 (1997), pp. 3367–3383. DOI: 10.1016/S0042-6989(97)00181-8.

[2] Horace B. Barlow. "Possible principles underlying the transformations of sensory messages". In: *Sensory Communication*. Ed. by Walter A Rosenblith. Vol. 6. 2. Cambridge, MA: MIT Press, 1961. Chap. 13, pp. 217–234. ISBN: 9780262518420. DOI: 10.7551/mitpress/9780262518420.003.0013. URL: http://www.trin.cam.ac.uk/horacebarlow/21.pdf.

[3] Bruce C Hansen and Edward A Essock. "A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes". In: *Journal of vision* 4.12 (2004), pp. 5–5.

[4] Joseph J Atick. "Could information theory provide an ecological theory of sensory processing?" In: *Network: Computation in Neural Systems* 3.2 (May 1992), pp. 213–251. ISSN: 0954-898X. DOI: 10.1088/0954-898X/3/2/009.

[5] V Billock. "Neural acclimation to 1/f spatial frequency spectra in natural images transduced by the human visual system". In: *Physica D: Nonlinear Phenomena* 137.3-4 (2000), pp. 379–391. ISSN: 01672789. DOI: 10.1016/S0167-2789(99)00197-9.

[6] Bruce C Hansen and Edward A Essock. "Influence of scale and orientation on the visual perception of natural scenes". In: *Visual Cognition* 12.6 (2005), pp. 1199–1234. ISSN: 1350-6285. DOI: 10.1080/13506280444000715.

[7] David J Tolhurst, Yoav Tadmor, and T. Chao. "Amplitude spectra of natural images." In: *Ophthalmic and Physiological Optics* 12.2 (1992), pp. 229–232. ISSN: 0275-5408. DOI: 10.1111/j.1475-1313.1992.tb00296.x.

[8] A. van der Schaaf and J.H. van Hateren. “Modelling the Power Spectra of Natural Images: Statistics and Information”. In: *Vision Research* 36.17 (1996), pp. 2759–2770. ISSN: 00426989. DOI: 10.1016/0042-6989(96)00002-8.

[9] David J. Field. “Relations between the statistics of natural images and the response properties of cortical cells”. In: *Journal of the Optical Society of America A* 4.12 (1987), pp. 2379–94. ISSN: 0740-3232. DOI: 10.1364/JOSAA.4.002379.

[10] G J Burton and Ian R Moorhead. “Color and spatial structure in natural scenes”. In: *Applied Optics* 26.1 (1987), pp. 157–170. ISSN: 0003-6935. DOI: 10.1364/AO.26.000157.

[11] Aude Oliva and Antonio Torralba. *Modelling the shape of the scene: A holistic representation of the spatial envelope*. 2001. DOI: 10.1023/A:1011139631724. URL: http://people.csail.mit.edu/torralba/code/spatialenvelope/%20http://www.springerlink.com/index/K62TG81W8352G71H.pdf (visited on 10/22/2011).

[12] Carlos Alejandro Párraga and David J. Tolhurst. “The effect of contrast randomisation on the discrimination of changes in the slopes of the amplitude spectra of natural scenes”. In: *Perception* 29.9 (2000), pp. 1101–1116. ISSN: 0301-0066. DOI: 10.1068/p2904.

[13] Aaron P Johnson et al. “The magnitude of center – surround facilitation in the discrimination of amplitude spectrum is dependent on the amplitude of the surround”. In: *Journal of Vision* 11.7 (2011), pp. 1–10. DOI: 10.1167/11.7.14.

[14] Carlos Alejandro Párraga, Tom Troscianko, and David J. Tolhurst. “The effects of amplitude-spectrum statistics on foveal and peripheral discrimination of changes in natural images, and a multi-resolution model.” In: *Vision research* 45.25-26 (2005), pp. 3145–68. ISSN: 0042-6989. DOI: 10.1016/j.visres.2005.08.006.

[15] Sarah L Elliott, Mark A Georgeson, and Michael A Webster. “Response normalization and blur adaptation: Data and multi-scale model”. In: *Journal of Vision* 11.2 (2011), pp. 1–18. DOI: 10.1167/11.2.7.

[16] Michael A. Webster, Mark A Georgeson, and Shernaaz M Webster. “Neural adjustments to image blur.” In: *Nature neuroscience* 5.9 (Sept. 2002), pp. 839–840. ISSN: 10976256. DOI: 10.1038/nn906. URL: http://www.ncbi.nlm.nih.gov/pubmed/12195427%20http://www.nature.com/doifinder/10.1038/nn906.

[17] Michael A. Webster and Eriko Miyahara. “Contrast adaptation and the spatial structure of natural images”. In: *Journal of the Optical Society of America A* 14.9 (1997), pp. 2355–2366. DOI: 10.1364/JOSAA.14.002355.

[18] Alan A. Stocker and Eero P. Simoncelli. “Noise characteristics and prior expectations in human visual speed perception”. In: *Nature Neuroscience* 9.4 (2006), pp. 578–585. ISSN: 10976256. DOI: 10.1038/nn1669.

[19] Ahna R Girshick, Michael S Landy, and Eero P Simoncelli. “Cardinal rules: visual orientation perception reflects knowledge of environmental statistics”. In: *Nature neuroscience* 14.7 (2011), pp. 926–932.

[20] Ruyuan Zhang, Oh-Sang Kwon, and Duje Tadin. “Illusory Movement of Stationary Stimuli in the Visual Periphery: Evidence for a Strong Centrifugal Prior in Motion Processing”. In: *Journal of Neuroscience* 33.10 (2013), pp. 4415–4423. DOI: 10.1523/JNEUROSCI.4744-12.2013.

[21] D C Knill, David J. Field, and D Kersten. “Human discrimination of fractal images.” In: *Journal of the Optical Society of America. A, Optics and image science* 7.6 (1990), pp. 1113–1123. ISSN: 0740-3232. DOI: 10.1364/JOSAA.7.001113.

[22] Yoav Tadmor and David J Tolhurst. “Discrimination of changes in the second-order statistics of natural and synthetic images”. In: *Vision Research* 34.4 (Feb. 1994), pp. 541–554. ISSN: 00426989. DOI: 10.1016/0042-6989(94)90167-8. URL: http://www.ncbi.nlm.nih.gov/pubmed/8303837%20http://linkinghub.elsevier.com/retrieve/pii/0042698994901678.

[23] David J Tolhurst and Yoav Tadmor. “Band-limited contrast in natural images explains the detectability of changes in the amplitude spectra”. In: *Vision research* 37.23 (Dec. 1997), pp. 3203–15. ISSN: 0042-6989. DOI: 10.1016/S0042-6989(97)00119-3.

[24] Bruce C. Hansen and Robert F. Hess. “Discrimination of amplitude spectrum slope in the fovea and parafovea and the local amplitude distributions of natural scene imagery”. In: *Journal of Vision* 6.7 (2006), pp. 696–711. ISSN: 1534-7362. DOI: 10.1167/6.7.3.

[25] Bruno Richard et al. “Spatial summation of broadband contrast”. In: *Journal of Vision* 19.5 (May 2019), p. 16. ISSN: 1534-7362. DOI: 10.1167/19.5.16.

[26] M. G. A. Thomson and David H Foster. “Role of second- and third-order statistics in the discriminability of natural images”. In: *Journal of the Optical Society of America A* 14.9 (Sept. 1997), pp. 2081–2092. ISSN: 1084-7529. DOI: 10.1364/JOSAA.14.002081. URL: http://www.opticsinfobase.org/abstract.cfm?URI=josaa-14-9-2081.

[27] Michael A. Webster. “Visual Adaptation”. In: *Annual Review of Vision Science* 1.1 (2015), pp. 547–567. DOI: 10.1146/annurev-vision-082114-035509.

[28] Michael A Webster. “Adaptation and visual coding.” In: *Journal of vision* 11.5 (2011), pp. 1–23. DOI: 10.1167/11.5.3.

[29] C. Blakemore and F W. Campbell. “On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images”. In: *The Journal of physiology* 203.1 (1969), pp. 237–260. ISSN: 00223751. DOI: 10.1113/jphysiol.1969.sp008862.

[30] Hugh R. Wilson and Richard Humanski. “Spatial frequency adaptation and contrast gain control”. In: *Vision Research* 33.8 (1993), pp. 1133–1149. DOI: 10.1016/0042-6989(93)90248-U.

[31] April Schweinhart, Patrick Shafto, and Edward A. Essock. “Distribution of content in recently-viewed scenes whitens perception”. In: *Journal of Vision* 17.3 (2017), p. 8. DOI: 10.1167/17.3.8.

[32] Peggy Series and Aaron Seitz. “Learning what to expect (in visual perception)”. In: *Frontiers in Human Neuroscience* 7 (2013), p. 668. ISSN: 1662-5161. DOI: 10.3389/fnhum.2013.00668. URL: https://www.frontiersin.org/article/10.3389/fnhum.2013.00668.

[33] Edward A Essock et al. “Oblique stimuli are seen best (not worst!) in naturalistic broad-band stimuli: A horizontal effect”. In: *Vision Research* 43.12 (2003), pp. 1329–1335. ISSN: 00426989. DOI: 10.1016/S0042-6989(03)00142-1.

[34] Bruce C Hansen et al. “A cortical locus for anisotropic overlay suppression of stimuli presented at fixation”. In: *Visual Neuroscience* 32 (2015), E023. ISSN: 0952-5238. DOI: 10.1017/S0952523815000255.

[35] Andrew M Haun and Edward A Essock. “Contrast sensitivity for oriented patterns in 1/f noise: contrast response and the horizontal effect.” In: *Journal of vision* 10 (2010), p. 1. ISSN: 1534-7362. DOI: 10.1167/10.10.1.

[36] H Akaike. “A new look at the statistical model identification”. In: *IEEE Transactions on Automatic Control* 19.6 (Dec. 1974), pp. 716–723. ISSN: 0018-9286. DOI: 10.1109/TAC.1974.1100705. URL: http://ieeexplore.ieee.org/document/1100705/.

[37] Brian J Fischer. “Bayesian estimates from heterogeneous population codes”. In: *The 2010 International Joint Conference on Neural Networks (IJCNN)*. 2010, pp. 1–7. DOI: 10.1109/IJCNN.2010.5596687.

[38] Wendy J Adams, Erich W Graf, and Marc O Ernst. “Experience can change the ’light-from-above’ prior.” In: *Nature neuroscience* 7.10 (2004), pp. 1057–1058. ISSN: 1097-6256. DOI: 10.1038/nn1312.

[39] Grigorios Sotiropoulos, Aaron R. Seitz, and Peggy Seriès. "Changing expectations about speed alters perceived motion direction". In: *Current Biology* 21.21 (2011), R883–R884. DOI: https://doi.org/10.1016/j.cub.2011.09.013.

[40] Wei Ji Ma. "Organizing probabilistic models of perception". In: *Trends in Cognitive Sciences* 16.10 (2012), pp. 511–518. ISSN: 1364-6613. DOI: https://doi.org/10.1016/j.tics.2012.08.010. URL: https://www.sciencedirect.com/science/article/pii/S136466131200201X.

[41] Alan A. Stocker and Eero P. Simoncelli. "Sensory Adaptation within a Bayesian Framework for Perception". In: *Proceedings of the 18th International Conference on Neural Information Processing Systems*. NIPS'05. Vancouver, British Columbia, Canada: MIT Press, 2006, pp. 1289–1296.

[42] Peng Zhang et al. "Effects of Orientation-Specific Visual Deprivation Induced with Altered Reality". In: *Current Biology* 19.22 (2009), pp. 1956–1960. DOI: https://doi.org/10.1016/j.cub.2009.10.018.

[43] Deep Ganguli and Eero P Simoncelli. "Efficient sensory encoding and Bayesian inference with heterogeneous neural populations". In: *Neural computation* 26.10 (2014), pp. 2103–2134.

[44] David H Brainard. "The psychophysics toolbox". In: *Spatial Vision* 10.4 (1997), pp. 433–436. ISSN: 0169-1015. DOI: 10.1163/156856897X00357.

[45] Denis G. Pelli. "The VideoToolbox software for visual psychophysics: Transforming numbers into movies". In: *Spatial Vision* 10.4 (1997), pp. 437–442. ISSN: 0169-1015. DOI: 10.1163/156856897X00366.

[46] Nicolaas Prins and Frederick A. A. Kingdom. *Palamedes: Matlab routines for analyzing psychophysical data.* 2009. URL: http://www.palamedestoolbox.org/.

[47] Christian Kaernbach. “Simple adaptive testing with the weighted up-down method”. In: *Perception & psychophysics* 49.3 (Mar. 1991), pp. 227–229. ISSN: 0031-5117. DOI: 10.3758/BF03214307. URL: http://www.ncbi.nlm.nih.gov/pubmed/2011460%20http://www.springerlink.com/index/10.3758/BF03214307.