Temporal aspects of orientation pooling using visual noise stimuli

Nicole D. Anderson
Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, Ontario, Canada

Kathryn M. Murphy
Schools of Biomedical Engineering, and Computational & Engineering Science, McMaster University, Hamilton, Ontario, Canada

David G. Jones
Departments of Electrical and Computer Engineering, and Psychology, Neuroscience, and Behaviour, Schools of Biomedical Engineering, and Computational & Engineering Science, McMaster University, Hamilton, Ontario, Canada

Previous psychophysical studies have shown that oriented signals are pooled over space to support better perceptual performance. We have investigated whether oriented signals may also be pooled over time. Using an orientation-in-noise paradigm, the threshold to discriminate large orientation differences was measured as the minimum amount of signal required to discriminate an oriented stimulus from unoriented noise. Discrimination thresholds were better with dynamic stimuli, containing multiple independent samples over time, than with static stimuli presented for an equal duration. Thresholds for dynamic stimuli showed gradual improvement from very brief (12 ms) to remarkably long presentation times (>4 s). Spatial integration of orientation signals is very efficient and can be understood in terms of a model based on signal detection theory, with performance limited by early and late stages of intrinsic noise. The nature of temporal integration is different, however, and is more consistent with probability summation of the outputs from low-level orientation detectors operating at a very brief timescale but whose outputs can be combined over a very long duration to yield better perceptual performance.

Keywords: orientation, noise, discrimination, temporal integration, computational model, probability summation, psychophysics


Introduction

Early stages of visual processing depend on orientation-selective responses of neurons in primary visual cortex. In computational models of visual processing, these receptive fields are most often modeled as a set of oriented band-pass filters (DeAngelis, Ohzawa, & Freeman, 1993; Wilson & Bergen, 1979) that must be integrated across space to extract signal from noise and yield a reliable response for perceptual decision making. The quality of the pooled response depends upon the number of samples: As more independent samples are combined, the signal-to-noise ratio of the pooled response improves. Many psychophysical studies of orientation discrimination have shown that there is spatial pooling of oriented filters because thresholds improve as the size of the stimulus increases. For example, orientation discrimination thresholds improve as a result of increasing the length of an oriented line or grating (Andrews, 1967b; Howell & Hess, 1978; Westheimer & Ley, 1997) or increasing the area of an orientation-in-noise stimulus (Jones, Anderson, & Murphy, 2003). Spatial integration of orientation information seems to be an obligatory function that operates even when pooling local orientation signals from a central target and surrounding distractors impairs performance on a task (Anderson, Murphy, & Jones, 2001; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). Taken together, these examples suggest that spatial pooling of orientation-selective filters is a fundamental component of the visual processing that underlies orientation discrimination.

Temporal pooling of neural responses can also contribute to better visual performance. Longer presentation times facilitate detection and discrimination of brightness and low-contrast stimuli (Breitmeyer & Ganz, 1977; Legge, 1978; Nachmias, 1967) and lead to better discrimination of fine changes in orientation tilt (Andrews, 1967a; Watt, 1987; Zlatkova, Vassilev, & Mitov, 2000). In contrast, we have previously found that discrimination of a large
orientation-in-noise difference does not improve with longer stimulus presentations (Jones et al., 2003). This distinction between large and small orientation discriminations may arise because thresholds for larger orientation differences operate on a shorter time course than thresholds for smaller orientation differences (Zlatkova et al., 2000). Our earlier findings suggested that the information required for optimal discrimination of large orientation differences was available within 100 ms. Increasing the stimulus presentation time beyond that did not provide additional information. Those studies, however, used static stimuli and only increased the stimulus duration rather than presenting more independent temporal samples of the orientation signal. If temporal pooling of orientation information operates in a manner similar to spatial pooling, then using dynamic stimuli should lead to discrimination thresholds that improve with increasing stimulus duration.

We have taken advantage of our recently developed orientation-in-noise paradigm (Jones et al., 2003) to study temporal pooling of orientation signals. The visual stimuli are high contrast, broadband in spatial frequency, and contain an oriented signal embedded in unoriented noise. The strength of the orientation signal can be varied from 0% to 100%, and the threshold is the smallest amount of orientation signal necessary to discriminate an oriented signal from unoriented noise. Temporal pooling was assessed by making the stimulus dynamic, using multiple independently generated frames of the orientation noise stimulus. Because the dynamic stimulus is broadband in temporal frequency, it will recruit both transient and sustained neural mechanisms. This dynamic stimulus effectively increases the number of samples in proportion to the duration of the stimulus and is ideal for testing whether temporal pooling contributes to the discrimination of large orientation differences.

In this study, we compared discrimination thresholds for static and dynamic orientation-in-noise stimuli. With increasing stimulus size, thresholds improved due to spatial pooling of orientation signals. We describe these results in terms of a model for orientation discrimination based on signal detection theory, where performance is limited by early and late stages of intrinsic noise. For a given stimulus size and equal stimulus durations, thresholds were always better with dynamic stimuli. Dynamic thresholds showed gradual improvement from very brief (12 ms) to remarkably long presentation times (>4 s), suggesting that performance may depend on higher level integration mechanisms. Whereas spatial pooling is very efficient and provides nearly optimal integration, temporal pooling is consistent with low-level orientation detectors operating at a very brief timescale, the outputs of which are combined using probability summation to yield better perceptual performance that continues to improve over a very long duration. Overall, these findings demonstrate that orientation signals are integrated over time, as well as space, to overcome noise in the stimulus and uncertainty in low-level neural mechanisms.

### Methods

#### Subjects

Three subjects participated in the experiments. One observer (N.A.) was highly practiced on the tasks; two observers (J.C. and P.B.) had no practice before testing and were naive as to the purpose of the experiment. All three subjects had normal or corrected-to-normal visual acuity.

#### Visual stimuli

We used two types of orientation-in-noise stimuli (Global and Local) that we have recently developed to measure orientation discrimination (Jones et al., 2003). Both stimuli were high-contrast gray-level patterns, containing a broad range of spatial frequencies, with a variable amount of orientation signal in the presence of unoriented noise. To generate the stimuli, pixel values were selected from a uniform random distribution of 256 gray levels. For both kinds of stimuli, the parameter \( P \) represents the strength of the orientation signal, namely, the percentage of pixels carrying the orientation signal versus noise pixels. When \( P = 100\% \), the stimulus was a one-dimensional white noise grating with either horizontal or vertical orientation; when \( P = 0\% \), the stimulus was uniform random noise, with no dominant orientation.

The Global orientation stimulus was generated by starting with a one-dimensional horizontal or vertical white noise grating, randomly selecting \((100 - P)\%\) of the pixels, and replacing them with new random gray levels (Figure 1). The term “Global” was used because the stimulus contained partially occluded oriented contours that extended across the entire stimulus. The Local orientation stimulus was generated by sequentially assigning gray levels to pixels along each orientation stripe. With probability \( P/100 \), the next pixel continued with the same gray level; otherwise, a new randomly chosen gray level was assigned. Each orientation stripe was drawn independently (Figure 1). The term “Local” was used because the stimulus contained oriented contour segments with independent gray levels and no inherent continuity extending across the entire stimulus.

The visual stimuli were generated on an Apple Macintosh G4 computer in the MATLAB (MathWorks) programming environment using Psychophysics Toolbox routines (Brainard, 1997). Stimuli were randomly generated before each trial, with dynamic stimuli created by presenting a new image on each frame or by assigning
new values to the lookup table on each frame. Both animation techniques were found to give equivalent thresholds. The resulting dynamic stimuli were broadband in temporal frequency content (Figure 2).

Stimuli were displayed on a 17-in. Apple Studio Display color monitor at a resolution of 800 × 600 pixels, a frame rate of 85 Hz, and a mean luminance of 40 cd/m². Stimuli were presented within a circular aperture. Seven aperture diameters were used: 0.7°, 2°, 3.5°, 5°, 7°, 10.5°, and 14°. All viewing was binocular from a distance of 57 cm. In the first experiment, static and dynamic stimuli were presented for 100 ms. In the second experiment, nine presentation times were used: 1, 2, 4, 9, 17, 43, 85, 340, and 1,360 frames. These correspond to presentation times of 11.8, 23.6, 47.2, 106, 201, 507, 1,000, 4,000, and 16,000 ms, respectively.

Procedure

Orientation discrimination thresholds were determined using a two-interval forced-choice (2IFC) paradigm. Subjects were instructed to judge which interval contained the oriented stimulus and to respond by pressing one of two keys. Within a trial, one interval contained a horizontally or vertically oriented stimulus with a variable amount of orientation signal, and the other interval contained an unoriented noise stimulus (0% orientation signal). A trial started with a central fixation point (500 ms) followed by the first interval, a uniform black screen (500 ms), and the second interval.

Threshold estimates for each experimental run were based on 60 trials, with the amount of orientation signal varied according to the QUEST method (Watson & Pelli, 1983). Thresholds were calculated at the 82% correct level, corresponding to the point of maximum slope on the underlying Weibull psychometric function (Harvey, 1986). Each experimental run consisted of two interleaved staircases to measure thresholds separately for the vertical and horizontal stimuli. No orientation biases were

Figure 1. Orientation-in-noise stimuli used in experiments. Stimuli are composed of an oriented signal component and an unoriented noise component. Examples are presented for the Global orientation stimulus (top) and the Local orientation stimulus (bottom). When the orientation signal strength is either 100% or 0%, the two kinds of stimuli are identical; for intermediate levels of orientation signal, they are dissimilar. Global stimuli contain oriented contours that extend across the entire stimulus, whereas Local stimuli contain local oriented contours. All of these examples contain a vertical orientation signal.

Figure 2. Movies illustrating static and dynamic versions of Global (up) and Local (down) orientation-in-noise stimuli. Each movie presents a static and dynamic version of the stimulus for both vertical and horizontal orientations. All stimuli contain 30% orientation signal. Click on each image to view the corresponding movie.
observed, and thus, mean thresholds were collapsed across orientation. Mean thresholds for each subject were based on six experimental runs.

Computational model

We formulated a computational model of local orientation processing based on pooling simple cell responses over space and time. The spatial pooling was implemented as ideal linear integration over a very brief duration (1 frame). The temporal pooling was implemented as probability summation. This model was implemented in two equivalent forms: as an analytical model by fitting theoretical equations to the threshold data and as a computer simulation by implementing the steps of the model in MATLAB to compute model thresholds for comparison with human thresholds.

Model analysis

To analyze spatial pooling, we plotted orientation discrimination thresholds as a function of aperture size and fit them with a theoretical equation (Equation 1, described later) that describes performance in terms of two stages of intrinsic noise. The “early noise” stage corresponds to the first stage of orientation processing and reflects uncertainty at the level of signal filtering. The “late noise” stage corresponds to uncertainty in the decision stage after the orientation signals have been pooled. Best fitting parameters were determined for the full set of psychophysical thresholds using a method of least squares.

To analyze temporal pooling, we plotted orientation discrimination thresholds as a function of stimulus duration on a log–log scale. According to signal detection theory, a slope of −0.5 would indicate an ideal integrator, but data lying along a line with a shallower slope are more suggestive of probability summation of noisy orientation detectors operating over a brief timescale. Assuming probability summation, the best fitting parameters (threshold and slope) of the psychometric functions of these low-level orientation detectors were determined from the threshold data.

Model simulation

This computer simulation extends a spatial pooling model described previously (Jones et al., 2003). The first stage is concerned with spatial integration of orientation signals. To simulate the perceptual decision made by human observers in a 2IFC task (with brief presentation time), we implemented a MATLAB function to indicate which of two stimulus images contained more orientation signal. Each stimulus image was convolved with oriented filters designed to resemble horizontal and vertical simple cell receptive fields. Three sizes of filter were used, with the smallest being $3 \times 3$ pixels ($[-1 \ 2 \ -1; -1 \ 2 \ -1; -1 \ 2 \ -1]$) and others being $6 \times 6$ and $9 \times 9$ versions of the same even symmetric kernel. (Similar results were obtained when using a quadrature pair of oriented filters to compute orientation energy.) This linear filter has an orientation bandwidth of ±27° and a spatial frequency bandwidth of 1.6 octaves and corresponds to a diameter of approximately 0.25°, 0.5°, or 0.75°, which is comparable to the receptive field properties found in the central visual field of macaque V1 (Hubel & Wiesel, 1968). For each image, within each orientation channel, the responses were full-wave rectified and pooled by summing over the entire area of the stimulus. The absolute value of the difference between the summed responses in the two orientation channels was taken as a measure of the orientation signal in the image. A vertical stimulus yielded a large vertical response and small horizontal response, whereas an unoriented noise stimulus yielded small responses in both channels and thus a difference near zero. Because of the noise in the stimulus, this procedure occasionally chose the wrong stimulus just as our human subjects did, especially when orientation signal and stimulus size were both small. The strength of the orientation signal in the stimulus was varied systematically (from 0% to 100%), and the percentage of correct responses from 10,000 trials was recorded to obtain the model psychometric function (Figure 7).

The second stage of the model is concerned with temporal pooling of orientation signals for dynamic stimuli with extended presentation times. To simulate probability summation, for each frame of the stimulus, a noisy decision was made regarding which of two alternative images was more oriented. The probability of choosing one image over the other was given by the psychometric function computed in the first stage of the simulation. At the end of the trial, the stimulus that was chosen the majority of times was selected as the response for that trial. This procedure was repeated for 10,000 trials to compute the proportion correct for a given orientation signal strength and stimulus duration. The orientation signal strength for which the model was 82% correct was taken as the model threshold (Figure 6, dashed lines).

Results

Orientation discrimination thresholds were first measured for three subjects using both static and dynamic versions of the Global and Local orientation noise stimuli with a 100-ms presentation time. In general, the thresholds were better for dynamic than for static stimuli (Figure 3). Thresholds were consistent within subjects, as indicated by the small error bars, and similar between subjects.
Thresholds were poorer at the smallest aperture size (0.7°) and improved with larger stimuli. Optimal thresholds for the Local stimuli (8.0% dynamic, 9.8% static) were achieved at an aperture size of 3.5° or larger. For the Global stimulus, thresholds improved rapidly when the aperture size was increased from 0.7° to 3.5°, followed by a slower rate of improvement for larger aperture sizes. This general pattern of threshold improvement with increasing aperture size is the same as we have previously reported and demonstrates spatial pooling of oriented signals (Jones et al., 2003).

Discrimination thresholds for the Local stimuli were considerably better than thresholds for the Global stimuli as we have previously reported (Jones et al., 2003). For a given orientation signal strength, the percentage of pixels carrying orientation information was the same in the Global and Local stimuli; however, the Local orientation stimulus is more effective because it is a better match for the location orientation processing mechanisms in the human visual system. For example, at a 2° aperture size, with a static Local stimulus, subjects require 14.8% orientation signal at threshold, whereas with a static Global stimulus, subjects required almost twice as much signal (27.1%) to discriminate the orientation from unoriented noise. Despite this difference, switching from static to dynamic led to a similar degree of improvement for both Local and Global stimuli (Figure 3). On average, at a given aperture size, subjects viewing the Global stimulus required 4.4 ± 0.6% less orientation signal at threshold when it was dynamic, and for the Local

![Figure 3. Orientation discrimination thresholds for three observers using static and dynamic stimuli. Thresholds for Global (left, circles) and Local (right, triangles) stimuli improved with increasing aperture size, indicating spatial integration. Thresholds also improved with dynamic stimuli (closed symbols) compared with static stimuli (open symbols), indicating temporal integration. Both static and dynamic stimuli were presented for 100 ms. Error bars represent ±1 SEM.](image-url)
stimulus, they required 3.7 ± 1.3% less orientation signal when it was dynamic.

Because providing more samples within the 100-ms dynamic stimuli led to improved orientation discrimination thresholds, we tested whether providing additional samples within 1 s would lead to an even greater improvement. We have shown previously that thresholds for 100-ms and 1-s presentation times are not different for static stimuli (Jones et al., 2003). In contrast, for dynamic stimuli, increasing the presentation time to 1 s led to better orientation discrimination thresholds (Figure 4). At the smallest aperture size (0.7°), approximately 10% less orientation signal was needed at threshold for the 1-s dynamic Global stimulus and approximately 7% less signal at threshold for the 1-s dynamic Local stimulus. It is interesting to note that the threshold for the smallest aperture size of the 1-s dynamic Local stimulus was comparable to that for the largest aperture size of the static Local stimulus, illustrating a space–time trade-off. Very low orientation discrimination thresholds can be obtained with either a large number of spatial samples or a large number of temporal samples.

Signal detection theory (Green & Swets, 1966) provides a method for expressing orientation discrimination thresholds for the Global and Local orientation-in-noise stimuli in terms of early and late components of intrinsic noise (Jones et al., 2003). In this framework, performance is predicted by the following equation:

\[
\text{threshold} = Z\sqrt{\sigma_E^2 + \frac{k^2\sigma_L^2}{N}},
\]

where \(k\) is a parameter related to the effectiveness of the type of orientation stimulus in evoking a neural response, \(\sigma_E\) is the standard deviation of the early noise, \(\sigma_L\) is the standard deviation of the late noise, \(N\) is the number of independent spatial samples and is proportional to the aperture size, and \(Z\) is the \(z\) score corresponding to the threshold measurement criterion. In the present experiments, thresholds were obtained at 82% correct, corresponding to a \(z\) score of 0.915. This analytical model was fit to the entire set of threshold data using the method of least squares (Figure 5). Performance was captured well with a common source of late noise assumed for all conditions (\(\sigma_L = 4.3\)). For both Global and Local stimuli, best fitting early noise parameters were lower for the dynamic than for the static stimuli and were lower for the 1-s dynamic stimuli than the 100-ms dynamic stimuli: \(k\sigma_E = 43.2\) (Global static), 35.3 (Global dynamic 100 ms), 22.9 (Global dynamic 1 s); \(k\sigma_E = 26.5\) (Local static), 17.9 (Local dynamic 100 ms), 7.8 (Local dynamic 1 s).
Performance was well described by this equation ($r^2 = .98$). It is not possible with our experimental data to tease apart the parameters $k$ and $A_E$. These results suggest that at the early stage of intrinsic noise, the overall amount of variability or uncertainty of the neural signal is lower with the dynamic stimuli.

Thresholds improved to approximately the same extent for Global and Local stimuli. The joint parameter $kA_E$ provides a description of the effectiveness of the oriented signal, with larger values corresponding to greater uncertainty about the orientation of the stimulus. For the Global stimulus, $kA_E$ improved by 7.9% orientation signal (from 43.2 to 35.3) with the 100-ms dynamic stimulus, whereas $kA_E$ improved by 8.6% orientation signal (from 26.5 to 17.9) with the Local 100-ms dynamic stimulus. Moreover, the degree of improvement was surprisingly consistent across aperture sizes. Thresholds were, on average, 1.28 times better (range = 1.17–1.38) for all sizes of the 100-ms dynamic Global stimulus and, on average, 1.31 times better (range = 1.21–1.51) for all sizes of the 100-ms dynamic Local stimulus. The similar improvements observed for all spatial conditions may reflect a fundamental mechanism of temporal integration that underlies performance with dynamic stimuli and operates independently of spatial pooling mechanisms.

In tilt discrimination tasks, temporal integration times are estimated to be on the order of between 50 and 100 ms (Watt, 1987; Zlatkova et al., 2000). Our results, however, indicate that oriented signals can be integrated over at least a second to promote better discrimination. This suggests that the temporal integration in the present orientation-in-noise experiment may reflect different mechanisms than those mediating temporal integration in a tilt discrimination task. To understand the relationship between presentation time and threshold improvement further, we explored the nature of the temporal integration using the smallest aperture ($0.7^\circ$) presented for different durations, ranging from very brief (12 ms) to very long (16 s). Thresholds for both Global and Local stimuli improved with increasing presentation time (Figure 6). Threshold improvement was shallow for the shortest presentation times, followed by a rapid improvement when the presentation time was increased beyond 47 ms. Thresholds continued to improve until the presentation time was increased beyond 4,000 ms for both Global and Local stimuli. Best fitting lines to the data showed that the rate of improvement was much shallower than $-0.5$ that would be predicted by a square root of $N$ rule. The slopes of the best fitting line for the Global and Local stimuli are $-0.14$ and $-0.22$, respectively. These slopes are more consistent with a probability summation account of threshold improvement, where improvement rates are typically shallower than the rate predicted by an ideal integrator (Graham, 1977; Watson, 1979).

Performance was well described by this equation ($r^2 = .98$). It is not possible with our experimental data to tease apart the parameters $k$ and $A_E$. These results suggest that at the early stage of intrinsic noise, the overall amount of variability or uncertainty of the neural signal is lower with the dynamic stimuli.

Thresholds improved to approximately the same extent for Global and Local stimuli. The joint parameter $kA_E$ provides a description of the “effectiveness” of the oriented signal, with larger values corresponding to greater uncertainty about the orientation of the stimulus. For the Global stimulus, $kA_E$ improved by 7.9% orientation signal (from 43.2 to 35.3) with the 100-ms dynamic stimulus, whereas $kA_E$ improved by 8.6% orientation signal (from 26.5 to 17.9) with the Local 100-ms dynamic stimulus. Moreover, the degree of improvement was surprisingly consistent across aperture sizes. Thresholds were, on average, 1.28 times better (range = 1.17–1.38) for all sizes of the 100-ms dynamic Global stimulus and, on average, 1.31 times better (range = 1.21–1.51) for all sizes of the 100-ms dynamic Local stimulus. The similar improvements observed for all spatial conditions may reflect a fundamental mechanism of temporal integration that underlies performance with dynamic stimuli and operates independently of spatial pooling mechanisms.

In tilt discrimination tasks, temporal integration times are estimated to be on the order of between 50 and 100 ms (Watt, 1987; Zlatkova et al., 2000). Our results, however, indicate that oriented signals can be integrated over at least a second to promote better discrimination. This suggests that the temporal integration in the present orientation-in-noise experiment may reflect different mechanisms than those mediating temporal integration in a tilt discrimination task. To understand the relationship between presentation time and threshold improvement further, we explored the nature of the temporal integration using the smallest aperture ($0.7^\circ$) presented for different durations, ranging from very brief (12 ms) to very long (16 s). Thresholds for both Global and Local stimuli improved with increasing presentation time (Figure 6). Threshold improvement was shallow for the shortest presentation times, followed by a rapid improvement when the presentation time was increased beyond 47 ms. Thresholds continued to improve until the presentation time was increased beyond 4,000 ms for both Global and Local stimuli. Best fitting lines to the data showed that the rate of improvement was much shallower than $-0.5$ that would be predicted by a square root of $N$ rule. The slopes of the best fitting line for the Global and Local stimuli are $-0.14$ and $-0.22$, respectively. These slopes are more consistent with a probability summation account of threshold improvement, where improvement rates are typically shallower than the rate predicted by an ideal integrator (Graham, 1977; Watson, 1979).
The process of probability summation assumes a low-level mechanism that makes a probabilistic decision about the orientation signal in the stimulus over a brief timescale (τ). By accumulating these guesses over the duration of the stimulus, perhaps by means of a higher level mechanism, the subject’s responses improve with longer presentation times. This low-level orientation mechanism can be described in terms of a two-alternative forced-choice Weibull function:

\[ \psi(x) = 1 - 0.5 \exp \left( -\left( \frac{x}{a_1} \right)^\beta \right), \]

(2)

where \( \psi(x) \) is the proportion of correct responses, \( x \) is the percentage of the orientation signal in the stimulus, \( a_1 \) is the orientation discrimination threshold (at 82% correct) after one temporal sample, and the exponent \( \beta \) is the slope of the psychometric function (at threshold). Although the two alternatives occurred sequentially in our experiment, it is convenient (and mathematically equivalent) to consider them as occurring simultaneously and being compared moment by moment. This low-level mechanism performs a discrimination between the orientation signals and indicates which interval contains the larger signal. When the stimulus is below threshold, this mechanism is still correct 50% of the time. According to probability summation, during a long presentation time, the subject observes \( N \) independent temporal samples of the stimulus and selects the interval chosen by the low-level discrimination mechanism a majority of times. Analyzing how thresholds improve with increasing temporal samples seems complicated because there are many different combinations that will add up to the correct majority decision. Therefore, it is useful to do the analysis in terms of detection.

The same low-level mechanism can also be described in terms of an equivalent detection mechanism operating on the difference between the orientation signals in the two intervals. Whenever this difference (over a brief timescale τ) is above threshold, the detector signals “yes”; otherwise, it remains silent. The corresponding yes/no Weibull function is

\[ \psi'(x) = 1 - \exp \left( -\left( \frac{x}{a_1} \right)^\beta \right), \]

(3)

where \( \psi'(x) \) is the probability of detecting an orientation difference after one temporal sample. Because the difference between the orientation signals has twice the variance, the slope of the function is doubled, \( \beta' = 2\beta \). According to probability summation, following \( N \) temporal samples of the stimulus, the subject fails to detect the orientation difference only if the low-level discrimination mechanism failed \( N \) times. This unique event occurs with probability \( \exp(-x/a_1)\beta^N \). Consequently, the predicted threshold based on probability summation is

\[ a_N = N^{-1/\beta} a_1, \]

(4)

where \( a_N \) is the improved detection threshold following \( N \) temporal samples (Nachmias, 1981). Therefore, on a log–log plot of threshold versus time (Figure 6), the slope of the best fitting line \( (-1/\beta) \) provides an estimate of the exponent of the underlying psychometric function. The threshold value at presentation time τ provides an estimate of \( a_1 \). Although the data do not allow us to determine τ with precision, it is clear from the graph in Figure 6 that it could be as brief as 1 or 2 video frames (τ = 12–24 ms). Putting these together, assuming probability summation, the analytical prediction for the parameters of the underlying psychometric functions are as follows: \( a_1 = 51.5, \beta = 3.6 \) for the Global orientation stimulus, and \( a_1 = 26.7, \beta = 2.3 \) for the Local orientation stimulus. These are in good agreement with the actual psychometric functions measured in human subjects, although the QUEST procedure is optimized to obtain quick and accurate threshold estimates, with less precision in slope estimates.

Model simulation

This model for spatial and temporal pooling of orientation information was also implemented in MATLAB as a computer simulation (see the Methods section). The first stage was concerned with spatial integration and orientation discrimination of a static (or briefly presented) stimulus. Because of the spatial pooling of orientation signals, the model thresholds improved with increasing stimulus size and would asymptotically approach zero if

![Figure 7. Predicted psychometric functions of mechanisms underlying performance in the orientation-in-noise task. The open symbols show the results of the model simulation used to predict orientation discrimination thresholds for the Global (circles) and Local (triangles) orientation stimuli for a stimulus size of 0.7° and a presentation time of 1 frame (12 ms). Each open symbol represents the proportion correct out of 10,000 simulated trials. The parameters of the best fitting Weibull functions (dashed lines) were as follows: \( a_1 = 47.91, \beta = 3.64 \) (Global); Local \( a_1 = 28.17, \beta = 2.25 \) (Local). These are in good agreement with the results from human subjects and consistent with the analytical predictions based on probability summation working backward from the thresholds versus time data in Figure 6.](image-url)
not limited by late intrinsic neural noise that follows spatial summation. Assuming a late noise parameter \(\sigma_L = 4.3\), the computer simulation using the smallest receptive field size (0.25°) replicates the theoretical prediction shown in Figure 5 (dashed curve) for the static condition.

For the model, as was the case for human subjects, orientation discrimination thresholds for the Local orientation stimulus were consistently better than for the Global orientation stimulus. For a given orientation signal strength, although the same percentage of pixels carry orientation information in the Local and Global stimuli, the Local orientation stimulus is more effective because it is a better match for the local orientation processing mechanisms in both the model and the human subjects. This can be further illustrated by considering a particular example in detail. Figure 7 shows the model psychometric functions obtained for the Global (\(\alpha_1 = 47.91, \beta = 3.64\)) and Local (\(\alpha_1 = 28.17, \beta = 2.25\)) orientation stimuli for a stimulus size of 0.7° and a presentation time of just 1 frame. Both the thresholds and slopes are a good match for the human results and the analytical model (see above).

These model psychometric functions (Figure 7) form the basis for understanding the second stage of the model, which is concerned with temporal pooling of orientation signals for dynamic stimuli and extended presentation times. To simulate probability summation, this psychometric function was used to make a probabilistic decision (independently for each frame) about which of the two alternatives contained more orientation signal. The stimulus that was chosen the majority of times during the trial was taken as the model’s response. This procedure was repeated for each trial to determine model thresholds for Global and Local stimuli with presentation times from 12 to 16,000 ms (Figure 6, dashed lines). The thresholds obtained from the computer simulation of probability summation are a good match for the human thresholds and also confirm the theoretical analysis linking the exponent of the psychometric function, \(\beta\), and the slope of the thresholds versus time data on a log–log plot.

Probability summation provides a good account of both the Global and Local dynamic results, despite the different spatial characteristics of the stimuli, and the difference in their absolute thresholds. The predictive power of the probability summation model is impressive: A very careful measurement of the psychometric functions at a presentation time of 12 ms (Figure 7) enables the accurate prediction of orientation discrimination thresholds at presentation times that are longer by up to three orders of magnitude (Figure 6, dashed lines).

### Discussion

The results can be summarized in terms of three main findings. First, orientation discrimination thresholds are better with dynamic versus static stimuli. Second, the improvement in thresholds is similar for both Global and Local stimuli and all aperture sizes, despite their different spatial characteristics. Third, thresholds continue to improve for a remarkably long time—more than 4 s. Taken together, these results show that increasing the number of independent samples over time leads to better orientation discrimination thresholds. This is consistent with the notion that the signal-to-noise ratio of the underlying neural mechanisms improves when oriented signals are pooled over time as well as space.

The orientation discrimination thresholds are well described using a model designed to account for early and late stages of intrinsic neural noise. The threshold limit for larger aperture sizes is captured by the late-noise parameter, which reflects uncertainty at the decision stage after oriented signals are pooled, and is independent of the number of spatial and temporal samples. The threshold data are well fit when a common source of late noise was used for all of the experimental conditions. Thresholds for smaller aperture sizes and brief presentation times are most influenced by the parameter describing performance in terms of early noise, which reflects uncertainty at the earliest stage of filtering. Early-noise values were lower for dynamic stimuli, suggesting that the uncertainty at the level of orientation filtering was reduced when more temporal samples were added. Uncertainty could be reduced either because the dynamic component of the stimuli provided a better match for the spatiotemporal receptive fields of the underlying neural mechanisms or because the effective number of samples for making a perceptual decision was increased. The results from the second experiment, in which thresholds gradually improved with increasing duration of the dynamic stimuli, favors the latter possibility, with temporal samples of the oriented signal being combined according to probability summation.

The improvement in orientation discrimination found with dynamic stimuli may seem surprising in light of previous studies showing that dynamic noise impairs contrast discrimination (Legge, Kersten, & Burgess, 1987) and letter identification (Gold, Bennett, & Sekuler, 1999) to a greater extent than static noise. This apparent difference, however, can be understood by considering that dynamic masking noise may effectively increase the strength of the masking and, as a result, more effectively mask the underlying static signal. In these previous studies, the target orientation signal remained static, whereas in the current experiments, the orientation signal was dynamic. If dynamic stimuli reduce uncertainty within underlying neural mechanisms, then the dynamic properties in the stimuli should be more salient. Taken together, these results suggest that dynamic aspects of a stimulus, whether signal or noise, are more effective components for perception.

The remarkably low orientation discrimination thresholds that we obtained under ideal spatiotemporal conditions reveal an exquisite sensitivity to orientation. Orientation can be discriminated when as few as 5% of the stimulus...
pixels carry the oriented signal. Similarly impressive thresholds (2–5%) have been measured for motion signal-in-noise tasks (Baker, Hess, & Zihl, 1991; Burr & Santoro, 2001; Newsome & Pare, 1988). These performance limits found using signal-in-noise stimuli for both orientation and motion may therefore reflect a common property of the neural circuits that underlie each of these perceptual tasks. Perhaps the limit is set by intrinsic neural noise that occurs in the later stages of perceptual decision making.

Using the dynamic orientation-in-noise stimulus, thresholds improved gradually over a very wide range of durations—from 12 ms to 4 s. This suggests that temporal integration of orientation signals is based on probability summation. In this framework, the outputs from low-level orientation detectors operating at a very brief timescale are summed over a much longer time to arrive at the perceptual decision. It is not possible from the present experiments to determine precisely where the time constant of low-level detectors ends and probability summation begins, but the results are consistent with it being as brief as 12–24 ms. This is generally consistent with the finding that the stimulus energy required to identify large orientation differences is available within 20–50 ms of stimulus onset (Zlatkova et al., 2000). Moreover, these numbers are compatible with the estimated integration times of 40–80 ms of V1 simple cells (Hawken, Shapley, & Grosof, 1996).

It is worth emphasizing that the nature of spatial and temporal integration of orientation signals is different. For briefly presented orientation-in-noise stimuli of increasing size, the human visual system acts as an ideal integrator (Jones et al., 2003). Additional spatial samples provide improved discrimination thresholds that follow a steep “square root of N” rule. This is possible because all of the information is present at once; however, combining these signals must involve neural connections that cover either a long distance within V1 or connections with higher visual areas. For dynamic orientation-in-noise stimuli presented with increasing duration, it does not seem possible for all of the temporal samples to be stored and later combined using ideal integration. Instead, the visual system seems to make use of low-level orientation detectors that effectively make the best guess possible within a brief temporal sample, and these momentary guesses are accumulated over time using probability summation. This gives rise to the shallower rate of improvement that is found as presentation time increases. Interestingly, the visual system seems able to continue this process of probability summation indefinitely or at least as long as the patience and concentration of our subjects would allow.

Conclusion

In summary, the visual system integrates input over both space and time, enhancing our ability to extract signal from noise. In our orientation-in-noise task, subjects were able to integrate information across several degrees or over several seconds to obtain improved discrimination thresholds. Furthermore, the extent of spatial and temporal integration is several orders of magnitude larger than the spatial and temporal extents of the individual-oriented receptive fields. Spatial integration of orientation signals is very efficient and can be understood in terms of a model based on signal detection theory, with performance limited by early and late stages of intrinsic noise. The nature of temporal integration is different, however, and is more consistent with probability summation, where the outputs of low-level orientation detectors operating at a very brief timescale are combined over a very long duration to yield better perceptual performance.

Acknowledgments

This research was supported by grants to D.G.J. and K.M.M. from the Natural Science and Engineering Research Council of Canada and the Canadian Institutes of Health Research.

Commercial relationships: none.
Corresponding author: David G. Jones.
Email: djones@mcmaster.ca.
Address: McMaster University, Electrical and Computer Engineering, 1280 Main St W, Hamilton, ON, Canada L8S 4K1.

Footnote

1 Current address: Dept. of Psychology, Grant MacEwan College, Edmonton, AB, Canada T5J 4S2.

References


assessed with limited-lifetime random dot stimuli. *Journal of Neuroscience, 11*, 454–461. [PubMed] [Article]


