Combining orientations to identify letters

Moon Hee Lee

149-12 15 Dr.
Whitestone, NY 11357
Moopri@aol.com

Stuyvesant High School
345 Chambers St.
New York, NY 10282

Professor Denis Pelli's Laboratory
Psychology and Neural Science
New York University
6 Washington Place
New York, NY 10003

Submitted December 10, 2001 to the Intel Science Talent Search.

M. H. Lee (2002) Combining orientations to identify letters. Intel Science Talent Search. Semifinalist. http://psych.nyu.edu/pelli/#intel

Summary

Visual research has revealed mechanisms in our brain called “channels” that detect patterns of light and dark bands. In our daily lives we mostly see objects like letters, not simple patterns, so there must be a higher level in our brain that integrates channels to recognize the object. By systematically masking letters with visual noise, I discovered that the mechanism used for letter identification is unselective across orientations.

Combining orientations to identify letters

Abstract

Visual channels detect specific frequencies or orientations of periodic visual patterns, such as gratings. However, when we look at the world, we rarely see gratings. Mostly, we see complex objects. Aside from being used to detect simple patterns, what role do channels play in identifying real life objects like letters? Since letters contain multiple orientations, there must be a mechanism in the visual pathway that integrates information across orientations. I systematically measured the observers' ability to identify letters and to detect gratings in the presence of visual noise at particular orientations. Results for letters and gratings were very different. The experiment revealed that the visual channel used to identify letters integrates equally across all orientations, whereas the channel used to detect a grating is tuned to the orientation of the grating. This is unlike the result for frequency tuning, where the same selectivity is found for letter identification and grating detection (Solomon & Pelli, 1994). This indicates that the channel used for letter identification is different than that used for grating detection, having the same frequency tuning, but no orientation tuning.

Introduction

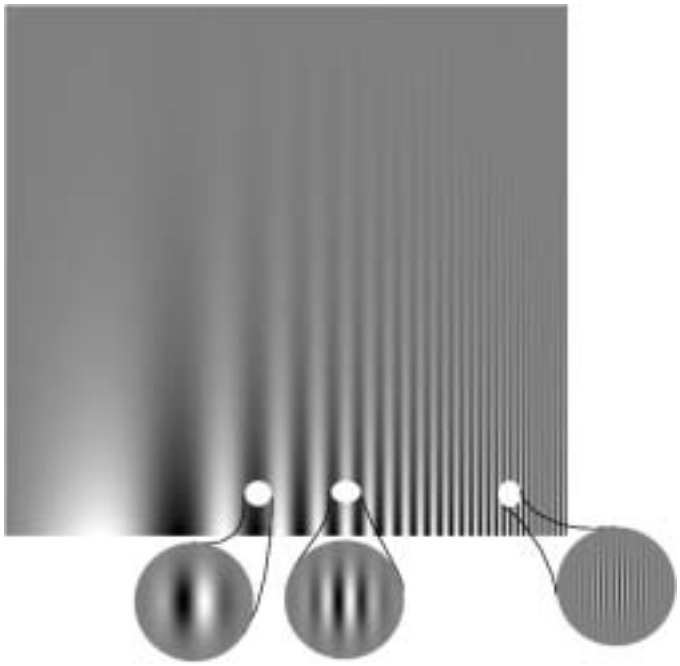
Identification of objects is essential to our daily lives. What each level of the visual system does to produce our visual experience has not been clearly elucidated in psychophysics or physiology. Although we do not see an object as an aggregate of spatial frequencies and orientations, these components induce response in the initial level of cortical processing in the brain (area V1). Each neuron has maximum sensitivity at a particular spatial frequency and orientation. Sensitivity decreases sharply as the given spatial frequency or orientation moves away from the preferred value. Gratings consisting of alternating light dark stripes have a dominant spatial frequency and orientation, but most objects of daily life do not have a dominant frequency or orientation. Instead, key objects like letters are spectrally broad, containing multiple

frequencies and orientations. Understanding how the simple responses of visual channels (or neurons in V1) are combined may be the key to understanding our normal visual experience.

GRATINGS AND CHANNELS

Figure 1 allows you to see your own visual sensitivity to gratings as a function of spatial frequency, showing that you are most sensitive to middle spatial frequencies. In general, any visual stimulus can be described as the sum of many gratings by using Fourier transformation. Fourier transformation represents any function as a sum of sine waves. For images, these sine wave gratings are characterized by their spatial frequency, contrast, orientation and spatial phase. Form perception in the initial level of cortical processing has been studied extensively using gratings. By measuring the contrast threshold of grating patterns over a wide range of spatial frequencies, Campbell and Robson (1968) showed there must be many detectors ("channels") within the visual system, each responding to a different range of spatial frequency. These channels work independently of one another. Each has its own unique preferred frequency and orientation.

1-a



1-b

FIGURE 1-a. Contrast Sensitivity Function. In the demonstration, the frequency of the grating increases from left to right, and its contrast decreases from bottom to top. By tracing out the upward limit at which you can still distinguish the bars, you are measuring your contrast sensitivity function. **FIGURE 1-b.** Sample gratings. The one on the left has the low spatial frequency and the one on the right has high spatial frequency.

ORIENTATION SPACE VS. FREQUENCY SPACE

Even though we do not experience an object as a sum of simple patterns, the first level of visual perception is performed by spatial-frequency and orientation selective channels that perform elementary detections. There must be a hierarchical structure in the visual cortex that integrates all the information from these simple detectors to finally perceive an object the way we do in real life. At the second level of integration, the responses from channels must be combined. Combination across spatial frequency and orientation has been studied by Olzak and Thomas (1998). They used an object consisting of two gratings. When the two gratings differed substantially in both spatial frequency and orientation, they found that observers could not integrate the two gratings. Instead, the integration of frequency and orientation are done independently of each other. They found that summation could occur over all frequencies within a limited orientation range, or could occur over all orientations with a limited frequency range.

PERCEIVING MULTIPLE FREQUENCIES OF LETTERS

Although much is known about the responses of visual channels, we still don’t know how these channels are selected when perceiving real objects. Real life objects are complex, containing a broad range of spatial frequency and orientation. They presumably stimulate many different channels. However, when Solomon and Pelli (1994) measured the channels for grating detection and letter identification, they found that observers use just one channel, the same channel, for both tasks. As shown in Figure 2, one particular noise frequency is most effective in masking letter identification.

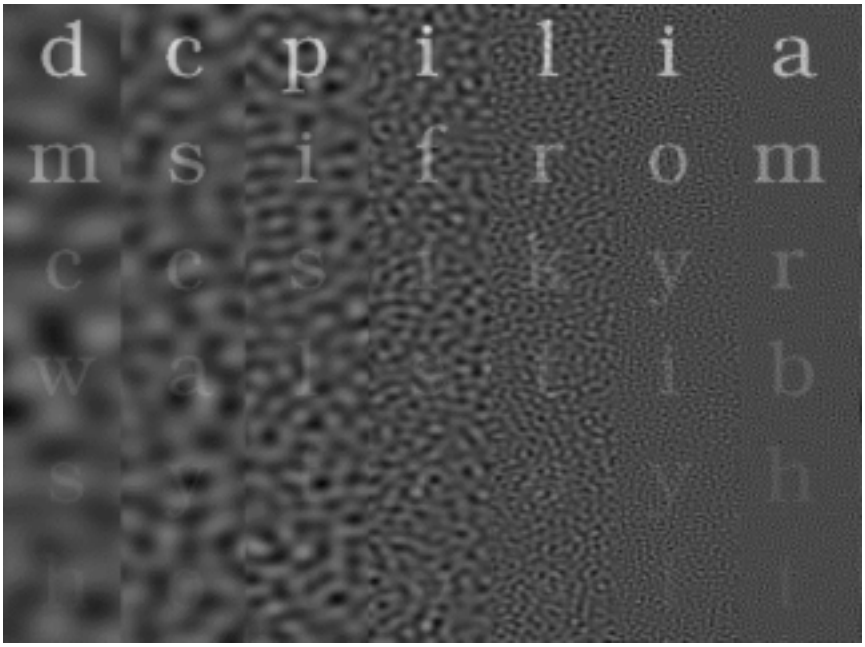


FIGURE 2. Letters in noise. Despite letters being broadband, noise at 3 c/letter (middle column) impedes letter identification more effectively than noise with lower or higher frequencies. Try it for yourself: Read down each column as far as you can; the positions of the faintest identifiable letters trace out the sensitivity of your eye's letter identification channel as a function of spatial frequency. (Solomon & Pelli, 1994).

WHAT ABOUT ORIENTATION?

Although it seems that only one frequency channel is used to identify letters, the observer's ability to integrate across many orientations is still a mystery. Gratings reveal the existance of frequency and orientation channels (Olzak & Thomas, 1999). On the other hand, broadband objects show that we do not integrate across frequencies (Solomon & Pelli, 1994). This still leaves open the question of how observers integrate across orientation, since letter identification presumably requires information from multiple orientations. Integrating features across orientation, the second level integrator might select multiple orientation channels, or just one channel tuned to an orientation. The focus of this research is to characterize how our visual system integrates across orientation to perceive a letter. The ultimate goal is to explain how we go from the elementary level of channels, detecting grating components, to perception of complex real life objects consisting of multiple spatial frequencies and orientations.

Methods and Procedure

STIMULI

Stimuli were created on Power Macintosh using MATLAB program and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Sloan letters, as specified by the NAS-NRC Committee on Vision (1980), were used for all signal letters. Sinewave gratings used for this experiment have the frequency of 1 c/deg and were displayed at an orientation of 120 degrees clockwise. All stimuli were displayed on a gamma-corrected computer monitor (Pelli and Zhang, 1983) with a video attenuator that drives just the green gun of the Apple 17" Multiscan color monitor. The viewing distance was 50 cm. The background luminance was 16 cd/m^2.

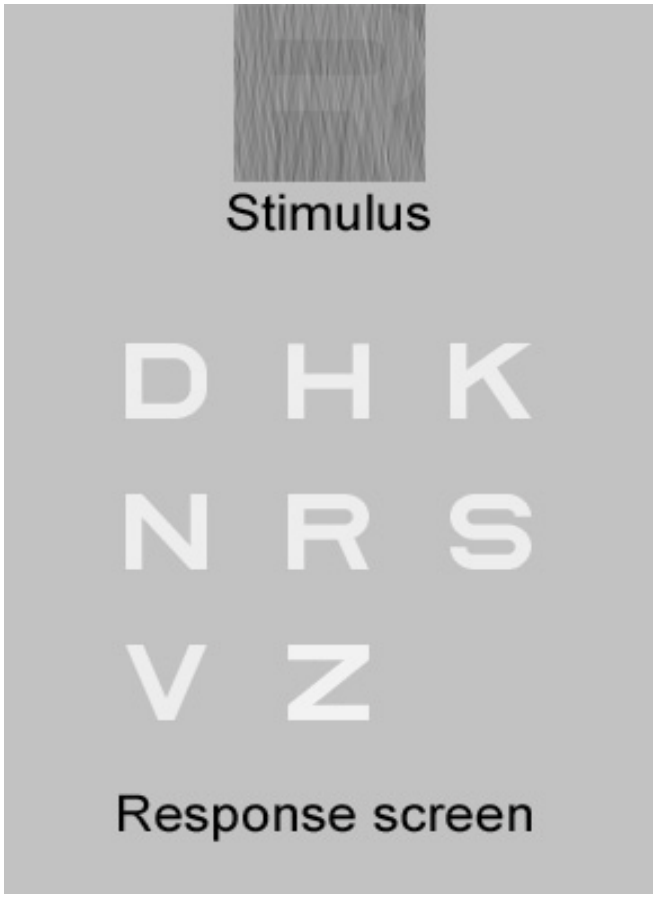


FIGURE 3. Letter Identification Screen. The signal letter is presented, briefly, in noise. The observer identifies the letter just seen by clicking the same letter in the response screen.

OBSERVER

Two observers with normal or corrected-to-normal vision voluntarily participated in the experiment. Based on Pelli, Burns, Farell, and Moore (2001), the observer's efficiency for identifying letters in a new font initially grows rapidly for the first 2,000 trials, and then rises much more slowly. Therefore, the first 2,000 trials of "learning period" were discarded.

LETTER IDENTIFICATION TASK

The observer stares at the fixation point before each trial. A random signal letter out of the possible eight, with noise, was presented at the fovea at the center of the screen for 200 ms. (See Fig. 3 top). The alphabet used for this task was: D, H, K, N, R, S, V, Z. After a 200 ms delay, the

observer was asked to identify the signal letter she saw from the eight possible signal letters, which were then displayed with 80% contrast. (See Fig. 3 bottom). When the observer correctly or incorrectly identified the signal letter, the contrast of the letter was lowered or increased, respectively. 40 trials per noise condition were used to measure the contrast threshold.

GRATING DETECTION TASK

The stimuli were again presented at the fovea. After the fixation point disappears, there were two 200 ms intervals of noise, separated by 200 ms of a blank screen. The grating was presented randomly in one of the two intervals. One interval had grating plus noise; the other had just noise. . (See Fig. 4.) The observer was then asked to click once or twice if she saw the grating in the first or the second interval, respectively. The threshold contrast was measured in the same way as for Letter Identification.

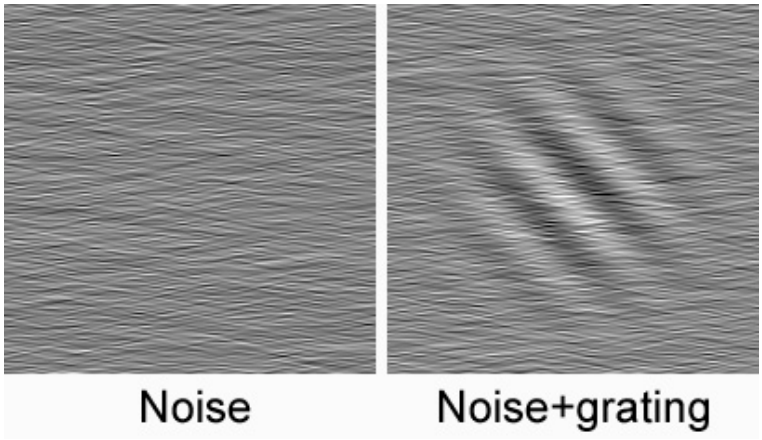


FIGURE 4. Grating Detection Screen. The two possible intervals for grating detection task are shown. Only one of the two intervals contains the grating. The observer is asked to identify which by clicking once or twice for the first interval or the second interval, respectively.

NOISE

To filter the noise orientation, a new noise filter was created. As shown in Figure 5, a wedge shaped filter passes noise at all frequencies but restricted to just a small band of orientation. Figure 5 shows the filter and the noise that results when the filter is applied to white noise. The filter is necessarily symmetric about the origin, because a visual image cannot contain imaginary components.

The bandpass noise was defined by a central orientation and its width. For example, a bandpass noise with central orientation of 50 degrees that is 20 degrees wide include the orientation from 40 to 60 degrees, and 220 to 240 degrees. We tested noise additivity by measuring the threshold elevation produced by each noise, with respect to the result obtained in

the blank field, to see if the sum of the elevation equals the elevation produced by the sum of the noise, i.e. white noise.

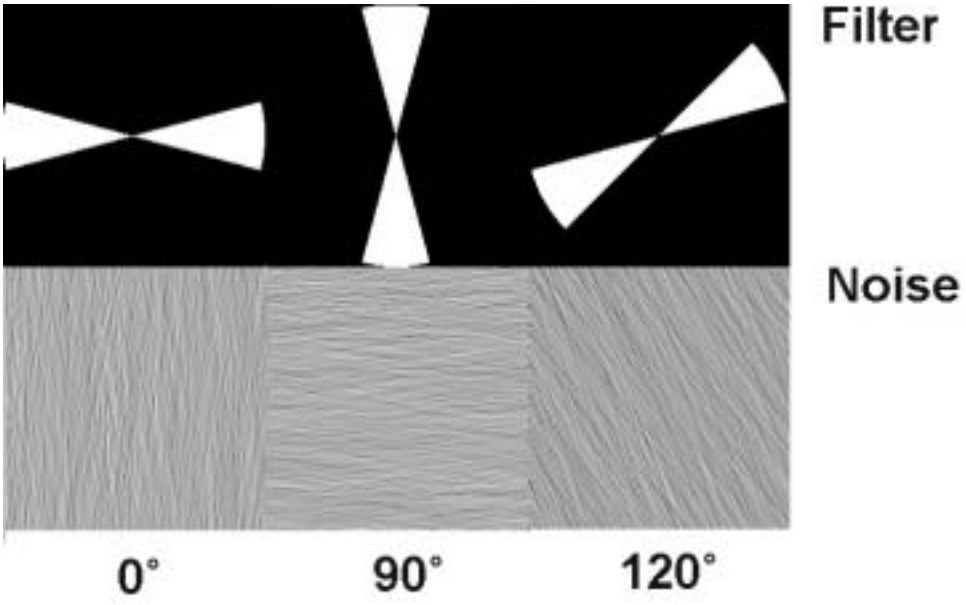


FIGURE 5. Filter and noise. The upper row shows the filter gain (white is 1, black is 0) as a function of x, y spatial frequency, with the origin in the middle. Each filter passes a 30 deg wide band of orientations. The second row shows samples of noise produced by applying the filter to white noise. The label below specifies the central orientation of the filter and noise.

THRESHOLD MEASUREMENT

The modified Quest staircase procedure (Watson and Pelli, 1983; King-Smith, Grigsby, Vingrys, Bones, and Supowit, 1994) was used to measure thresholds.

Results

By using the critical band masking technique, I measured the tuning of the visual mechanism that is used for performing these two tasks: letter identification and grating detection. Bandpass filters were used to systematically filter the noise. For example, a band that has a width of 30 deg, with the central orientation of 45 deg, would allow all orientations of the noise to pass through except for orientations from 30 to 60 deg and the reflection of that about the origin. I measured the contrast threshold for detecting a grating or identifying a letter as a function of the center frequency of the noise. I also measured the contrast thresholds without any noise (blank field) and with noise containing all possible orientations (white noise). A test of “noise additivity” was done to see if the sum of threshold elevation produced by several noises separately equals the threshold elevation produced by the sum of the noises.

CRITICAL BAND MASKING

Figure 6 shows the contrast threshold for letter identification as a function of central orientation for filters with bandwidth 20, 30, and 60 deg. The contrast elevation is increased as the bandwidth increases, because there is more noise. Across all orientations, the contrast threshold reached was 0.017 for 20 deg wide band; 0.020 for 30 deg wide band; and 0.027 for 60 deg wide band. The contrast threshold on a blank field is 0.014 and this threshold rises to 0.063 in the presence of white noise.

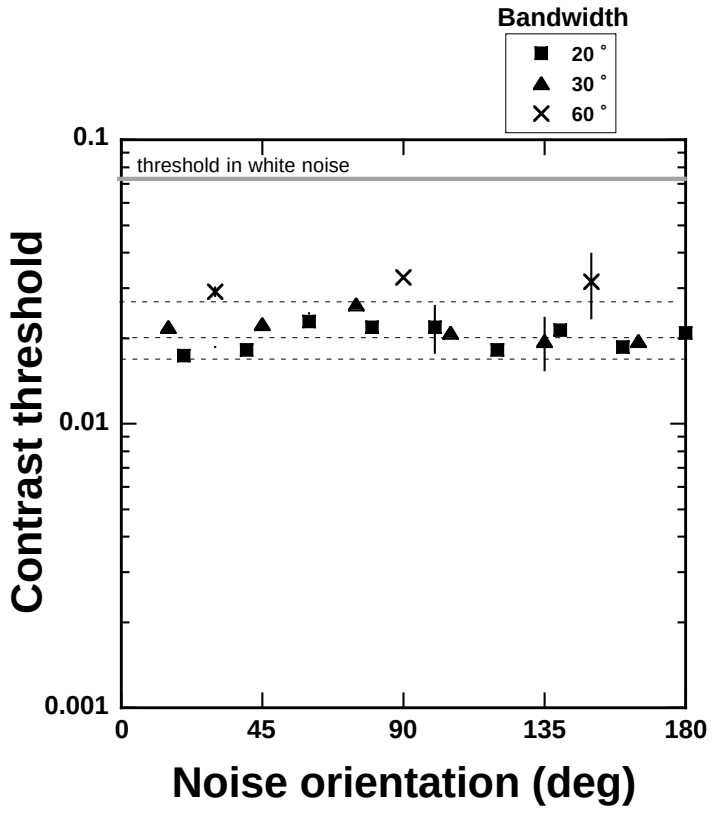


FIGURE 6. Contrast threshold as a function of noise orientation for letter identification. The threshold for bandwidth of 20, 30, and 60 deg is represented. There is no selective threshold elevation for any particular noise orientation. The dashed lines represent the additivity test prediction for one big channel that integrates across all orientations. This is calculated for each bandwidth used. All the data fall close to these flat, horizontal lines.

The contrast thresholds for each trial were again plotted with the grating detection with bandwidth of 30 deg (See Fig. 7). Gratings are detected by the channel tuned to the grating's orientation since a grating contains only a single orientation. The bandwidth of the channel at half the height of the channel is 75 deg. With letter identification, the observers' threshold elevation is flat, showing no selectivity at all for orientation.

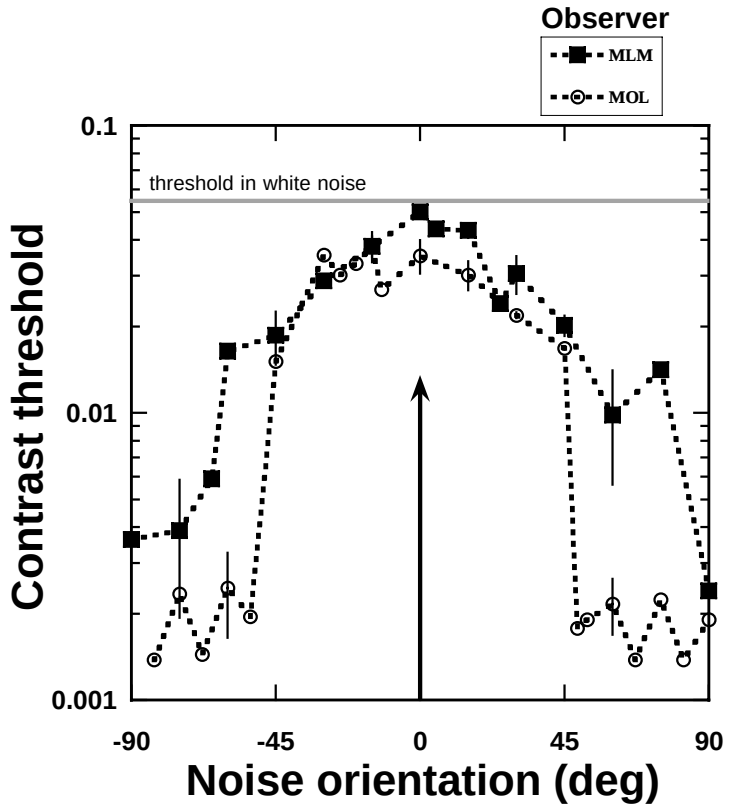


FIGURE 7. Contrast threshold vs. noise orientation for the grating detection task. Threshold is elevated over a narrow range of noise orientation, close to that of the grating. Noise at far orientations produces very little threshold elevation. This channel is narrowly tuned.

NOISE ADDITIVITY

The threshold elevation was measured by subtracting no-noise threshold energy from the threshold energy needed for each noise frequency. Then the threshold elevations were summed and compared with the threshold elevation in white noise, which is what one would get if one summed the noises themselves. Because the sum of the threshold elevations across noise bands is nearly equal to the threshold elevation white noise (i.e. the sum of the noises), there is noise additivity across orientation in letter identification (Fig. 8). Thus the thresholds rise linearly as we add more noise at various orientations. This implies that the noise is passing through the same channel.

Bandwidth (deg)	Sum of the threshold energies	Threshold energy in white noise
20	-1.15	-1.31
30	-1.20	-1.31
60	-1.25	-1.31

FIGURE 8. Noise Additivity Table. The sum of the threshold energies of the individual noise bands is nearly equal to the threshold elevation of the sum of the noises (i.e. white noise).

Discussion

Solomon and Pelli (1994) discovered that, although they expected spectrally broad letters to stimulate many channels, each tuned to a narrow range of frequencies, the visual system seems to ignore all but one channel, the one tuned to 3 c/letter. Further research in the frequency space by Majaj et al. (2001) concluded that the observer uses one channel and cannot learn to use multiple channels while doing the task—letter identification task is bottom up, and not top down as many people believed (Majaj, Pelli, Kurshan, and Palomares, 2001). It was also asserted that observer cannot integrate across frequencies (Solomon and Pelli, 1994, Majaj, Pelli, Kurshan, and Palomares, 2001). In principle, spectrally broad letters, not only induce the response of many channels tuned to a narrow range of frequencies, but many channels tuned to a narrow range of orientations respond as well. Therefore, we must also consider the mechanism used to integrate the multiple orientation presented by letters.

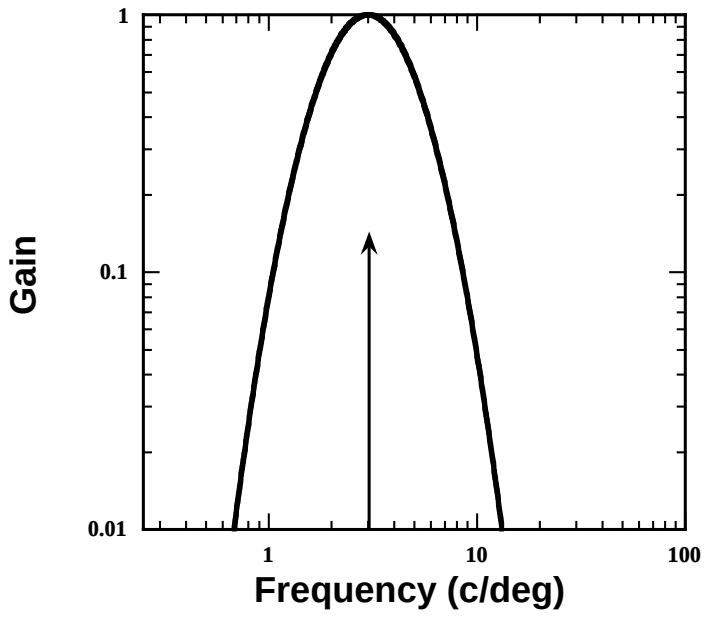


FIGURE 9. The tuning function of the spatial-frequency channel used for grating detection and letter identification (Solomon and Pelli, 1994).

Integrating across orientation seems to be unlike integrating across frequency. Solomon and Pelli found that to identify a letter, which contains multiple frequencies, we use one narrow channel similar to the one we use to detect a grating, which is a narrow band stimulus. (see Figs. 2 and 9). However, I found that the threshold elevation as a function of orientation is unselective (flat) for letter identification and tuned for grating detection. The noise additivity test showed that, with broadband objects, we use one broad channel integrating across all orientations. My results also show the use of an orientation tuned channel with narrow band gratings. Thus

different channels are used to identify letters (unselective in orientation) and to detect gratings (selective in orientation).

Comparison of the mechanism graphed in Fig. 6 for letter identification differ significantly from the selectively tuned channel presented in Fig. 7 for grating detection. Gratings contain only one orientation, 120 deg clockwise. Unlike with letter identification task, all orientations do not have the same effect on performance. Instead with grating detection, masking certain orientations hinders the reader from doing the task far more than masking other orientations. The channel used for perceiving one orientation, therefore, involve a narrowly tuned filter that is centered at one specific orientation. However, the mechanism used for letter identification, which is unselective across all orientations, differ from the one used for grating detection.

Critical band masking results for letter identification can be interpreted in two possible ways. The mechanism can consist of multiple separate channels that are tuned to specific orientation, like the one used for grating detection, working independently of each other. The other possibility is that the mechanism is really a large filter that integrates across multiple orientation channels. The noise additivity test, which can distinguish between the two alternatives, showed that the latter was true. The sum of the threshold elevations produced by the various noise bands was approximately equal to the threshold elevation produced by white noise (i.e. the sum of the noises). This is strong evidence that we have found a single channel that integrates linearly across all orientations, rather than a nonlinear combination of the activity of multiple channels.

The success of noise additivity across orientation is shown in Fig. 8. By taking the threshold elevation in white noise and dividing it by the number of bands, I was able to predict the threshold elevation produced by a narrower band, as indicated by the horizontal lines near the data points. Since the thresholds for multiple bands obey noise additivity, the flat tuning results show a single channel integrating across all orientations.

Conclusion

One might expect a spectrally broad letter to stimulate many frequency channels, but the observer is confined to the use of only one of them. It was expected that the second level integrator for combining multiple orientations for letter identification would be similar. My results show that the observer uses a single channel, as expected, but that it is unselective with respect to orientation, integrating all equally, even though the first level channel for grating detection is orientation tuned.

Acknowledgements

I would like to express my gratitude to Dr. Deis Pelli and Dr. Marialuisa Martelli for guidance, knowledge, insight, and advice. You have made my lab experience an enjoyable as well as an exciting one. I am grateful for the encouragement and many helpful comments that I received from all the members of Dr. Denis Pelli's Lab. I would especially like to acknowledge Isabelle Chen for teaching me how to organize my thoughts. I would also like to express my appreciation to Brian Pasley, Albert Leung, and Allen Ingling for assistance in programming my visual filter. I thank Dr. Flip Phillips for showing me how much fun science could be. Last but not least, thank you my wonderful teachers, Ms. Anne Manwell, Mr. Steve Farbstein, and Dr. John Utting, for encouraging me to do independent research and supporting me throughout the process.

References

Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat Vis, 10* (4), 433-436.

Campbell F.W., and Robson J.G. (1968) The application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197,* 551-566.

King-Smith, P.E., Grigsby, S. S., Vingrys, A.J., Benes, S.C., and Supowit, A. (1994) Efficiency and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation. *Vision Research, 34,* 885-912.

Majaj, N.J., Pelli, D.G., Kurshan, P., & Palomares, M. (2001). The role of spatial frequency channels in letter identification. *Vision Research,* in press.

National Academy of Sciences-National Research Committee on Vision (1980). Report of working group 39: Recommended standard procedures for the clinical measurement and specification of visual acuity. *Advances in Ophthalmology, 41,* 103-148.

Olzak, L. A., and Thomas, J. P. (1999) Neural recording in human pattern vision: model and mechanisms *Vision Research,* 39, 231-256.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10,* 437-442.

Pelli, D.G., Burns, C.W., Farell, B., & Moore, D.C. (2001). Identifying letters. *Vision Research,* in press.

Pelli, D. G., and Zhang, L. (1991) Accurate control of contrast on microcomputer displays. *Vision Research, 31,* 1337-1350.

Solomon, J. A., and Pelli, D. G. (1994) The visual filter mediating letter identification. *Nature, 369,* 395-397.

Watson, A. B., and Pelli, D. G. (1983) QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33,* 113-120.