

Computational Vision in Neural and Machine Systems, ed. L. Harris and M. Jenkin.  
Published by Cambridge University Press.  
c Cambridge University Press 2007.

These are the final proofs of the Graham & Wolfson chapter.  
Actual page numbers in the book are 9-47.

## **2. Exploring Contrast-Controlled Adaptation Processes in Human Vision (with help from *Buffy the Vampire Slayer*)**

**Norma Graham and  
S. Sabina Wolfson**

We have been interested for many years in intermediate levels of visual processing: levels which are lower than the perception of “objects” and “scenes” but higher than the pointwise processing of the retina and LGN<sup>1</sup>. Many of these intermediate processes are concerned with the initial analyses of pattern and form. Much of their action can be well modeled by what are technically linear operations – in particular, multiple analyzers sensitive to different ranges of spatial frequency and orientation<sup>2</sup>. However, some of their action cannot be modeled this way as it is fundamentally nonlinear. We have recently become interested in the dynamics of these intermediate nonlinear processes, and more particularly in questions about how the visual system sets its sensitivity based on the recent history of stimulation.

This chapter affords us an opportunity to be informal and to relate past work to present work in ways that are uncommon in journal papers. We are happy to take advantage of this opportunity. We will use informal speech and explanations and also personal anecdotes. And we will give many fewer references to published literature than is our wont, but instead will try to guide the reader to places where such references

---

<sup>1</sup>The terms “lower” and “higher” are only approximate, of course, since information travels “downstream” as well as “upstream.” Lennie (1998) presents interesting hypotheses – and an overall view – about the function and nature of the processes that have physiological substrates from V1 up to V4 and MT.

<sup>2</sup>The psychophysical research on these multiple analyzers, and a small amount of the physiological research, is described in Graham (1989) and summarized in Graham (1992).

can be found.

Two sets of psychophysical experiments – and the models that were tested by their results – are described in this chapter. The first is described briefly and the second at length.

The first set was designed to investigate behaviorally the dynamics of luminance-controlled processes like light adaptation in the retina or LGN. Strictly speaking, these processes are lower than the level that we have been most interested in and were done with a third major collaborator, Don Hood, who is very interested in that level. Further, this set is already published for the most part. Thus we will describe it quite briefly. However, we do describe it because it both inspired the second set and also gave us distinct expectations about how the second set would turn out.

The second set of experiments was designed to investigate the dynamics of contrast-controlled processes. We started out to study one such process that had proved necessary to explain our previous results with textured patterns (done in collaboration with other investigators, in particular Jacob Beck and Anne Sutter). But the results of this second set of experiments ended up suggesting the existence of an entirely different contrast-controlled process, and one that we had not previously even imagined. This second set of experiments and the new process they suggested will be the focus of most of this chapter.

## **2.1 Dynamics of luminance-controlled adaptation processes (light adaptation)**

### **2.1.1 Flickering the luminance of spatially-homogeneous backgrounds and measuring thresholds for superimposed luminance probes**

The first set of experiments, that we will only discuss briefly here, was intended to investigate the dynamics of luminance-controlled adaptation processes (e.g. light adaptation in the retina). Fig. 2.1 shows the spatial and temporal characteristics of this paradigm, which is often called the probed-sinewave paradigm.

In probed-sinewave experiments, the luminance of a spatially-homogeneous background is flickered sinusoidally in time during each trial. At some point during the trial a luminance-defined probe is introduced (of intensity  $\Delta I$ , an increment in the figure, but decrements have been used as well). It is typically a smaller disk in the middle of the flickering background.

You can see movies of these stimuli on the compact disk accompanying this book. Video 1 shows the flickering background disk by itself. Video 2 shows the flickering background disk with a probe increment introduced.

Results from one typical observer from one study are shown in Fig. 2.2 (with separate frequencies of flickering background in separate panels) and then again in Fig. 2.3 (with the results at different frequencies superimposed in one panel). Probe threshold is plotted as a function of phase, and, to help show trends in results, the phases are plotted through two cycles on the horizontal axis. The results in Figs. 2.2 and 2.3 show typical

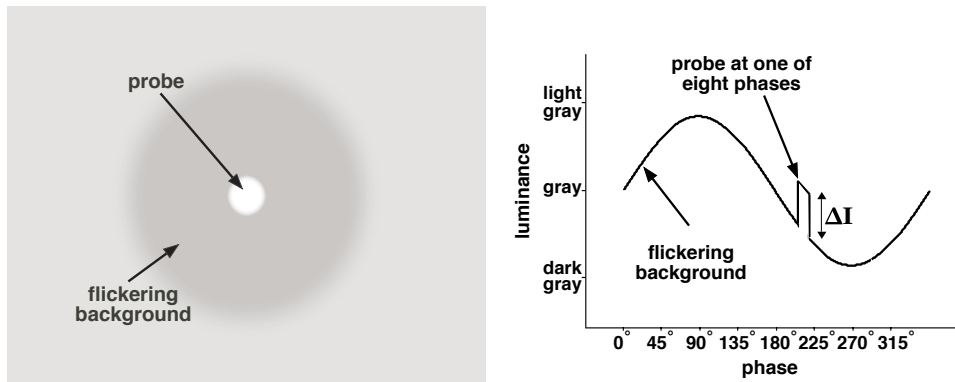


Figure 2.1: The probed-sinewave paradigm used to study luminance-controlled adaptation (light adaptation) processes. Spatial paradigm on left, and temporal paradigm on right. The luminance in the probe  $\Delta I$  is adjusted until the observer can just discriminate between the background-alone and the background-plus-probe.

features of experimental results from this paradigm. Of particular importance to the story here, note the large general increase in probe threshold magnitude (the upward displacement of the curves) as the background's flicker frequency is raised from lower (lighter and thinner lines and symbols) to higher (darker and bigger lines and symbols). The probe thresholds decrease at still higher frequencies – not shown here – until they are back down on average to the same level as at very low frequencies.

Many other studies – using different conditions and different observers – have been done by various groups of investigators. While the studies differed among themselves in a number of ways, they all showed a big general increase in probe threshold magnitude as the frequency of the flickering background increased from low to middling. (Many were compared in Graham, Wolfson, and Chowdhury, 2001.)

### 2.1.2 What do the results imply for models of light adaptation?

This empirical result – the general increase in probe threshold with increase in background flicker frequency – has turned out to be very powerful in discriminating among different models, or, more generally, in discriminating among different ideas of how light adaptation might work. Indeed Hood et al. (1997) showed that this empirical result completely rules out a large class of previously-successful models containing the best features of the two earlier modeling traditions (the merged models of Graham and Hood, 1992). This empirical result could not, however, immediately rule out in the same dramatic way a new model suggested by Wilson (1997). The Wilson model, based on explicit physiological pieces, could be trivially modified to do a satisfactory job to at least a good first approximation (Hood and Graham, 1998, as was subsequently shown also with a fuller set of results by Wolfson and Graham 2000, 2001a). But the Wilson model has some drawbacks (see discussion in Wolfson and Graham, 2001a,b). The best current candidate in our opinion is the more abstract model of

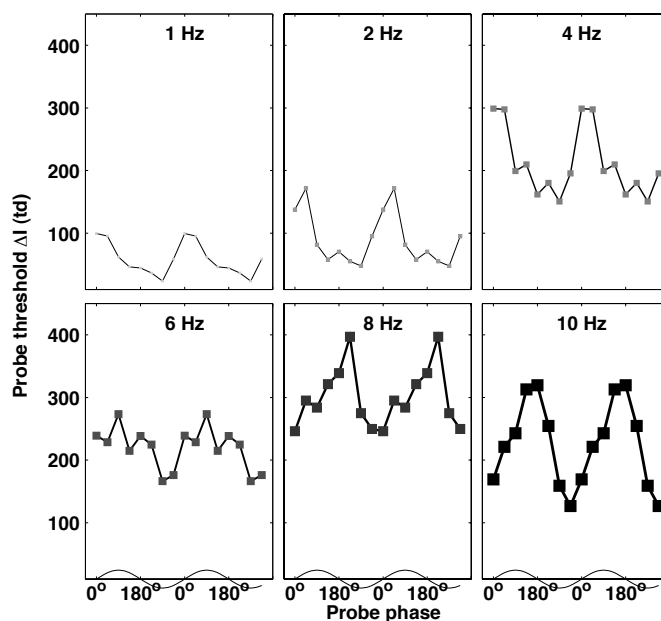


Figure 2.2: Experimental results from the probed-sinewave paradigm, plotted as probe threshold  $\Delta I$  versus phase. Two cycles of the background are shown for help in displaying the patterns in the results but the points in the second cycle are identical to those in the first cycle. Experimental data from observer JG in Fig. 4 of Hood, Graham, Wiegand, and Chase (1997).

Snippe, Poot, and van Hateren (2000, 2004) diagramed in Fig. 2.4. The Snippe et al. model with its three general kinds of processes easily handles the empirical result we have been talking about, that is, the general elevation of probe threshold as background flicker frequency increases from low to middling. The model does so by adding a contrast-gain-control process to the previously-suggested subtractive and divisive stages of luminance-controlled processes (light adaptation). See the figure legend for some more details. (The actual processes themselves are described precisely in Snippe et al., 2000.)

The contrast-gain control process in the model of Fig. 2.4 presumably acts before any stage at which there is substantial binocular combination, and therefore its physiological substrate is likely to be in the retina or LGN. This presumption comes from a further empirical result: In the probed-sinewave paradigm, most of the general elevation with increasing flicker frequency does not show interocular transfer (Wolfson and Graham, 2001b).

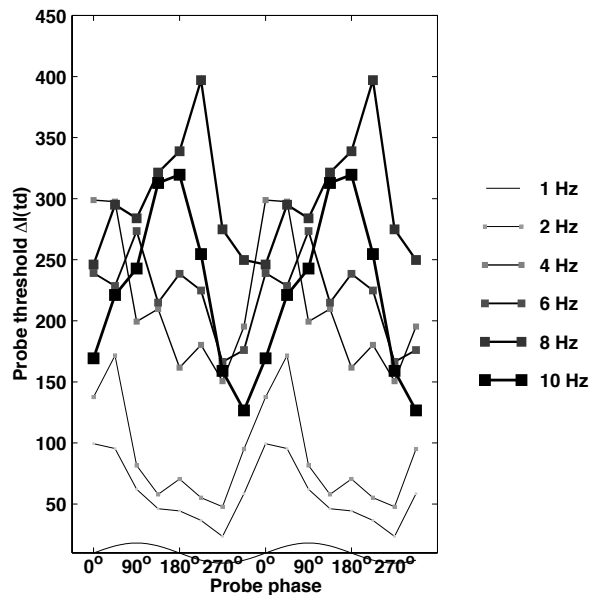


Figure 2.3: The experimental results for different frequencies in Fig. 2.2 superimposed in one panel. In general, the probe thresholds tend to increase as frequency of the flickering background is varied from low (lighter-colored and thinner symbols and lines) to middling (darker and thicker symbols and lines).

### 2.1.3 About some terms: additive or subtractive, multiplicative or divisive, and gain control

The Snippe et al. model (Fig. 2.4) has distinct processes labeled “divisive” and “subtractive.” This labeling distinction is quite suggestive and often useful. It sometimes causes difficulty, however. Consider the following: If a process simply multiplies two values (its inputs) to produce its output, what happens if you now redefine the inputs and outputs to be the logarithms of the originals? The process now adds the two newly-defined inputs to be its newly-defined output. Any difference between the old and new case is simply a matter of exactly what you say “input” and “output” mean. In the situations where the distinction is useful, the process is more complicated and/or the outputs and inputs do not lend themselves naturally to being redefined as logs (or antilogs). In general, it pays to be cautious about these terms and take them merely as suggestive.

The contrast-controlled process in the Snippe et al. model (Fig. 2.4) is called a “contrast gain control”, and is more of a multiplicative (or divisive) than a subtractive (or additive) process. Most of the processes for which people use the words “gain control” seem to have such a multiplicative nature, and the word “gain” often refers to

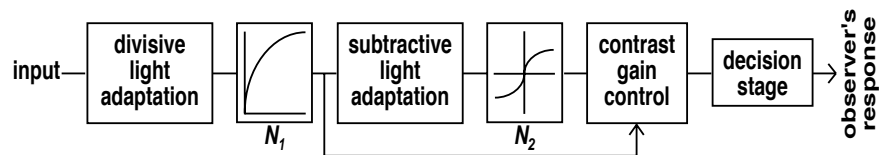


Figure 2.4: A diagram of the model of Snippe, Poot, and van Hateren (2000, 2004). This is the model for light adaptation that we currently find the most promising. They used three main processes because (in their words, p. 451, Snippe et al., 2000): “From psychophysical experiments on light adaptation (mainly using background steps), it has been concluded that light adaptation contains both divisive (also referred to as multiplicative) and subtractive components (e.g. Hayhoe et al., 1992; Graham & Hood, 1992). Our model follows this tradition. It has been suggested previously that the elevation of test thresholds on modulated backgrounds above the test detection level for a steady background could arise from a contrast gain control process (Hood et al., 1997; Wu et al., 1997). This gain control would be activated by the temporal contrast of the background modulation, which would decrease the transmission gain for the test pulse, and hence its detectability. The third module in our model, contrast gain control, is a quantitative implementation of this suggestion.” Precise description and equations for the separate processes are given in their articles. This diagram here is modified slightly from theirs (Snippe et al., 2000, Fig. 2) by some changes in notation and making further explicit the decision stage that leads to an observer’s response. We added the decision-stage box to emphasize the fact that the assumptions relating the outputs of processes relatively early in visual processing to an observer’s response are critical in the testing of any such model. Often quite straightforward assumptions prove reasonable and useful, however, and such is the case here.

a multiplicative constant relating the ratio of output to input.

We will use the words “luminance-controlled” and “contrast-controlled” adaptation processes to mean that either the luminance or contrast, respectively, of preceding visual stimulation is changing the responsiveness to new stimulation. By using these more general terms, we want to imply that the control at issue is not necessarily of the multiplicative sort that would be implied by “luminance gain control” or “contrast gain control.”

## 2.2 Dynamics of contrast-controlled adaptation processes

### 2.2.1 Flickering the contrast of patterned backgrounds and measuring the thresholds of superimposed patterned probes

In the second set of experiments – intended to investigate the dynamics of contrast-controlled adaptation processes – the contrast of a spatially-patterned background is

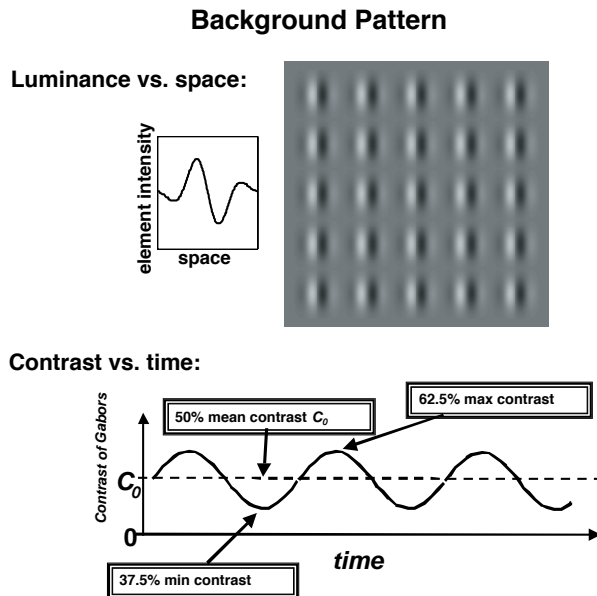


Figure 2.5: Diagram of the spatially-patterned flickering background stimulus used in our contrast adaptation experiments. The upper panel shows a piece of the background pattern with only 5x5 of the 15x15 Gabor-patch elements used in the experiment. The small inset on the left shows the Gabor function that describes the spatial profile of an individual element. The functions at the bottom show the temporal profile – contrast as a function of time – of each Gabor-patch element.

flickered sinusoidally in time during each trial. At some point during the trial a spatially-patterned probe is introduced.

You can see movies of these stimuli on the compact disk accompanying this book. Video 3 shows the flickering patterned background by itself. Video 4 shows the background with a probe introduced. The next few paragraphs and figures describe these stimuli further.

Figure 2.5 diagrams the general spatial and temporal characteristics of the flickering background in this kind of experiment. The upper part of Fig. 2.5 shows a piece of pattern. (This piece contains an array of 5x5 Gabor patches, whereas we used 15x15 in the experiments and in the movies.) The lower part of Fig. 2.5 shows the contrast in any one of these Gabor patches as a function of time. The values of the flickering background's contrast at the mean, peak, and trough that were used in the experiments reported here are labeled on the figure.

Figure 2.6 illustrates the presentation of a short-duration probe stimulus. The right-hand panel of Fig. 2.6 shows a piece of the background with probe at the moment the probe is presented, in other words, a test stimulus. To prevent confusion, it is important

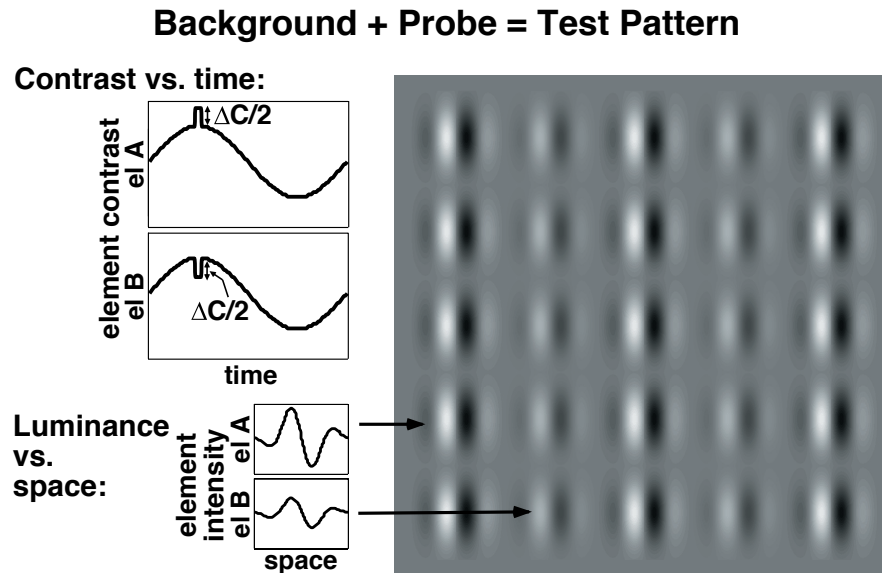


Figure 2.6: Example of the background-plus-probe or test pattern. The right panel shows a piece of the test pattern. The two types of Gabor patches that define alternating columns are called *element A* and *element B*, which are abbreviated *el A* and *el B*. The spatial profiles of these two element types are in the panels at the lower left showing luminance as a function of spatial position. The temporal profiles of the two element types are plotted in the panels at the upper left showing contrast as a function of time. The probe is an increase in contrast for *el A* and a decrease for *el B*. The increase and decrease are of equal magnitude  $\Delta C/2$ , so the total contrast difference between the two element types in the probe is  $\Delta C$ . Note that the word “test” here explicitly means the combination of background and probe. The word “probe” refers to the change that is made in the background pattern in order to produce the test pattern. In some contexts, either word can be used to the same effect, but in others (particularly later in this chapter) it will be useful to keep the meanings distinct.

to note that we will use the word “test” here to explicitly mean the *combination of background and probe*. The word “probe” refers to the change that is made to the background pattern. Thus, the test stimulus is the sum of the background stimulus and the probe stimulus. In Fig. 2.6 the illustrated test stimulus (the probe plus background) looks like alternating columns (vertical stripes) of different-contrast Gabor elements. Test stimuli can also be alternating rows (horizontal stripes) of Gabor patches (as in Fig. 2.7 left panel). We will call the two types of Gabor patches that define alternating rows or columns by the names *element A* and *element B*, abbreviated as *el A* and *el B* in Fig. 2.6. The luminance of element A and element B as a function of spatial position

## Observer's task: Horizontal vs. Vertical

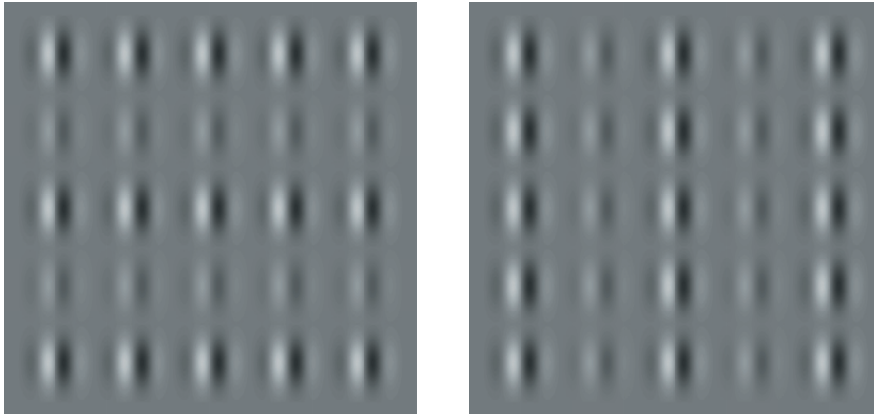


Figure 2.7: The observer's task is to identify the orientation of the contrast-defined stripes as being either horizontal (left panel) or vertical (right panel).  $\Delta C$  is varied to find the value at which the observer can just discriminate at criterion level between the two orientations. That value of  $\Delta C$  is called the probe threshold.

is plotted in the lower left drawings of Fig. 2.6.

The temporal characteristics of the probe presentation are shown in the upper left of Fig. 2.6. Contrast is plotted as a function of time for the two element types. The probe illustrated in Fig. 2.6 is composed of an increase in contrast for one type of element (*el A* in this case) and an equal-sized decrease in contrast for the other element. The total contrast in the probe will be called  $\Delta C$ , and so the contrast increase (*el A*) and decrease (*el B*) are each equal to half that quantity ( $\Delta C/2$ ) as indicated in the figure. (A forewarning: This symmetry between increase and decrease in contrast was true for all probes used in the experiments we are about to describe. However, later in the chapter, this will not be true, as we will explicitly point out at the appropriate point.)

The observer's task in our experiments reported here was always to identify the orientation of the contrast-defined stripes as being either vertical or horizontal (see Fig. 2.7) and, of course, these possibilities were randomly presented from trial to trial.

**Further details of the experiments reported here (for the interested reader).** On half the trials, the orientation of all the individual Gabor patches in a pattern was vertical as in the examples here. On the other half of the trials (randomly intermixed) the Gabor patches were all horizontal. Note that there is undoubtedly nonlinear distortion in all the illustrations here (the movies and the figures) but in the actual experiments, the stimuli were very carefully calibrated. For the experiments reported below, the spa-

tial frequency of the Gabor patches was 2 c/deg, and the other spatial attributes of the pattern were as pictured here. The background pattern flickered for a total of several seconds. The probe was presented approximately in the middle of the flickering period as dictated by the following constraints: There was always at least one second of flicker before and after the probe, and the flicker always started and stopped at positive zero-crossings. The probe could occur at 1 of 8 equally-spaced phases (randomly chosen on each trial) ranging from  $0^\circ$  to  $315^\circ$ . Between the periods of flicker the screen was a homogeneous field at the mean luminance of 58 cd/m<sup>2</sup>. The probe duration was 82 msec (7 frames at the CRT's refresh rate of 85 Hz) for observers JW and EG, and 35 msec (3 frames) for observers KF and SH. The observer responded to indicate whether the contrast-defined stripes were perceived as horizontal or vertical (see Fig. 2.7). The probe's contrast levels were determined on each trial using simple 1-up-3-down staircases. The probe threshold was calculated as the value of  $\Delta C$  such that the observer can just discriminate (at a criterion of 81% correct) between the two orientations. Observers were instructed not to fixate rigidly in order to avoid after-images. They were given feedback about correctness of their responses.

### 2.2.2 Why we used these patterns and what we expected to find

We used this type of pattern in these experiments because we already knew a great deal about their perception by human observers. What we knew made it sensible to believe the following, which we will briefly state and then elaborate upon.

The flickering background pattern (e.g., Fig. 2.5) should drive a particular contrast-gain-control process.

But it should NOT stimulate the pathway detecting the probe stimulus.

Thus, when we varied the temporal frequency of the flickering background and the phase of the probe with respect to this background and we then measured the observer's sensitivity to the probe, we hoped for the following outcome. We hoped that we would be measuring the dynamic properties of the contrast-gain-control process and that we would NOT be measuring (to any substantial extent) the dynamic properties of the pathway detecting the probe.

The following text, and the illustrations in Figs. 2.8 and 2.9, give a little more substance to the general idea just sketched.

The test pattern formed by the background-and-probe (see image in Fig. 2.6) is an example of the *element-arrangement textures* introduced by Jacob Beck and his colleagues in the early 1980's (e.g. Beck, Prazdny and Rosenfeld, 1983) and studied by many people, including us, since then. These textures are defined by the difference between two types of elements, in our experiments here the different levels of contrast in otherwise identical Gabor-patch elements. The particular arrangement of the two types of elements is striped in the example here, although other arrangements – in particular checked – have been used.

Note also that Gabor-patch element-arrangement textures like those used here (where contrast is the only difference between element types, e.g., Fig. 2.6) are very very similar to stimuli usually called *contrast-modulated noise* or *contrast-modulated sinusoidal gratings*. The spatial frequency and orientation of the Gabor patches in the element-arrangement textures play the role of the carrier spatial frequency and orientation in the

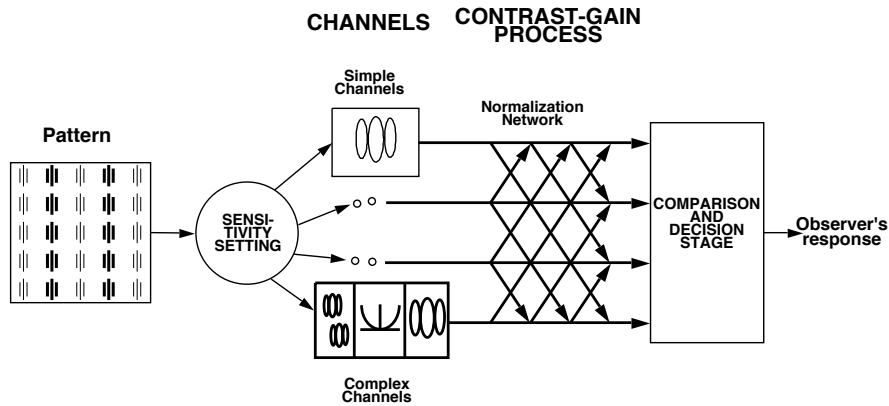


Figure 2.8: Our general model as it existed before the experiments described in this chapter. This framework includes complex channels and a contrast-gain control of the normalization type. See text for more details about these nonlinear processes. The “sensitivity setting” stage shown in this figure preceding the simple and complex channels would, in general, contain early local nonlinearities like retinal and LGN light adaptation. But we have shown that for experiments using texture element-arrangement patterns within the range of contrasts and luminances used here, this stage is effectively linear.

contrast-modulated noise or sinusoidal gratings. The spatial frequency and orientation of the stripes defined by the contrast differences in the element-arrangement textures play the role of the modulation signal in the contrast-modulated noise or sinusoidal gratings. Thus many results using contrast-modulated noise and sinusoidal gratings can be used also to support the statements we are about to make.

Thus one can say the following. A wide variety of experimental results using element-arrangement textures and other similar patterns can be explained on the basis of *spatial-frequency and orientation-selective channels* (which are linear filters), but – and it is a big caveat – only if at least two, dramatically different, nonlinear processes are also included in the model. We will describe these nonlinear processes in the next several paragraphs. We will not list all the many references for the statements we are going to make, but the interested reader can find the omitted references in the following publications: Graham and Sutter (1998, 2000), Graham and Wolfson (2001), Landy and Graham (2003), and Wolfson and Graham (2005a). These list many references to the studies that provide evidence for the existence of these nonlinear processes, their possible functions, and their general properties, as investigated by us and by many others.

As well as *spatial-frequency and orientation-selective channels* that are linear filters (which we call *simple channels* in Fig. 2.8 and which are sometimes called *first-*

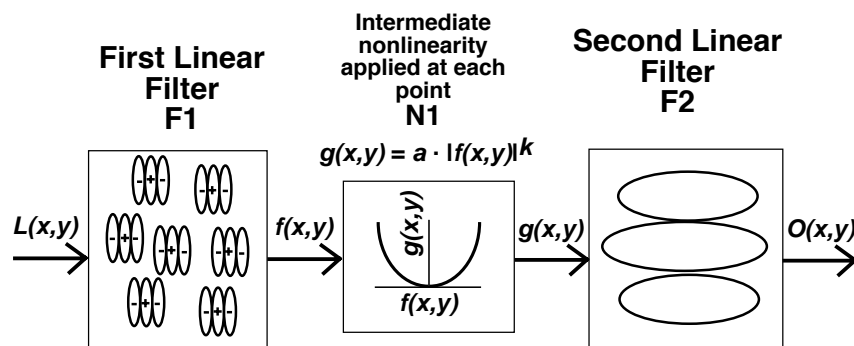


Figure 2.9: More details of the structure of a complex channel. Later in the chapter we will refer to this as the “original” kind of complex channel.  $L(x,y)$  is luminance at position  $(x,y)$ ;  $O(x,y)$  is the output of the complex channel;  $f$  and  $g$  represent functions specifying the outputs at intermediate stages within the channel; and  $a$  and  $k$  are parameters specifying the exact form of the intermediate nonlinearity  $N1$ .

*order channels*), it is necessary to postulate a nonlinear kind of channel. We call these nonlinear channels *complex channels*, and they are also sometimes called *second-order channels* or *non-Fourier channels*. As sketched in the bottom middle of Fig. 2.8 and drawn with details in Fig. 2.9, these complex channels contain two layers of linear filtering  $F1$  and  $F2$  separated by a nonlinearity  $N1$ . This kind of structure is often called, for short, an  $FNF$  structure. The nonlinearity sandwiched between the two linear filters is something like a full-wave or half-wave rectification applied to the output of the first filter at each spatial position and at each moment in time. (In particular, the output of  $N1$  is never less than zero.) The first linear filter  $F1$  is characterized by a relatively small receptive field (spatial weighting function) and the second linear filter  $F2$  is characterized by a substantially larger one.

In the case of element-arrangement textures made with Gabor patches, a simple linear channel would not be able to detect the striped element arrangement at all, but a complex channel would. We do not have the space here to go through this argument, but an interested reader can find a demonstration of this argument with an element-arrangement texture in Fig. 11 of Graham, Beck, and Sutter (1992) and with another kind of pattern in Fig. 4 of Landy and Graham (2003).

In addition to the rectification-type nonlinear process in the complex channels, we had discovered there was another and very dramatic nonlinearity. It depended upon the contrast of the patterns, and it was already very compressive for very low contrasts. We showed eventually that this compressive effect can definitely NOT be explained by any relatively-early local process occurring before the level of the channels. In particular, therefore, it cannot be explained by a luminance-controlled process like light

adaptation. We also showed that a contrast-gain-control process of the type frequently called *normalization* could explain this dramatic compressive effect. This normalization process involves inhibition among channels. Any given channel is inhibited by other channels in its *normalization pool*, where this inhibition acts in such a way (divisively) that it “normalizes” the response of the inhibited channel with respect to the total response from a rather large set of channels.

A normalization process of this sort can be shown to have a number of attractive functional properties that might explain why evolution produced such a process in the visual system. For example, it prevents overload on higher levels while it allows the channels to preserve their sensitivity and also their selectivity (e.g. for orientation) over a wide range of contrasts by changing the effective operating range on the contrast dimension. More recently it has also been suggested that such a process has the appropriate characteristics to help encode natural images efficiently.

For our element-arrangement patterns, with appropriate simplifying assumptions, we have previously derived a simple equation giving the predictions from the general model shown in Fig. 2.8. This equation – incorporating the normalization process as well as simple and complex channels with very few parameters – produces quantitative predictions that can account for a large body of results (e.g. Graham and Sutter, 2000).

Now let's use the general framework of Fig. 2.8 to translate the general expectations given at the beginning of this section into more specific statements about what we expected to find from the experiments using flickering patterned backgrounds.

First note that the observer's ability to identify the orientation of the element-arrangement test stimulus (e.g. Figs. 2.6 and 2.7) is mediated by the appropriately-tuned complex channels. By “appropriately-tuned” we mean the complex channels that have first-stage receptive fields matched in spatial frequency and orientation to the Gabor-patch elements, and second-stage receptive fields matched in spatial frequency to the striped arrangement (and sensitive to either vertical or horizontal orientation).

The flickering background (Fig. 2.5) produces NO response in these appropriately-tuned complex channels that can identify the test stimulus orientation because there is no striped arrangement in the background stimulus, and thus nothing for the second filter of these complex channels to respond to.

But the flickering background does drive the signal from the normalization pool up and down because it stimulates many simple channels and may also stimulate other complex channels that respond to the Gabor patches although they cannot identify the orientation of the striped arrangement.

Hence any effect of the flickering background on the response to the probe would be via its effect on the contrast-gain normalization process and not contaminated (to any large extent) by dynamic characteristics of the direct response to the probe.

On the assumption that the measured effects would be reflecting the characteristics of a contrast-gain control pathway of something like the normalization type, we planned to compare quantitatively the results of these experiments to predictions from a number of proposed models in the literature or constructed by us. The models included: the Carandini, Heeger, and Movshon (1997) model of normalization in V1 cells; the Wilson and Humanski (1993) model of a contrast-gain control based on some psychophysical results; a model consisting of a general purpose contrast-gain control

stage something like that in the Snippe et al. (2000) model of light adaptation (Fig. 2.4 here).

Although we had not done extensive quantitative predictions at the time we started collecting results from the experiments on flickering patterned backgrounds reported below, we knew quite a good deal about the general form of the predictions based on the above expectations.

In general there would be phase-specific effects and the precise form of them would change with temporal frequency of the background, reflecting the general dynamics of the normalization pool's inhibiting effect on the channel detecting the probe.

Some of the models would predict that, in spite of these phase-specific effects that depend on temporal frequency, the probe thresholds averaged across phase would not change very much with temporal frequency.

Other models, however, would predict a general elevation effect like that found in the earlier light adaptation results (Figs. 2.2 and 2.3) where the thresholds would generally elevate as temporal frequency increased from low to middling. Indeed it was exactly that effect that the Snippe et al. (2000) contrast-gain-control process (see Fig. 2.4) was designed to predict. And thus if we took their process as a model of the contrast-gain, the general probe-threshold elevation with background frequency would also be predicted for the contrast adaptation experiments.

### **2.2.3 The results on flickering patterned backgrounds**

Of course it took us much longer than we planned to finally get usable results from these contrast adaptation experiments when we did start trying to run them. The usual sorts of delays intervened including equipment breakdowns, software mistakes, and trouble getting the stimulus parameters in a range to produce measurable performance from typical observers (performance greater than chance and less than perfect). But this time we had another delay with nobody to blame but ourselves: when we finally collected usable results, we did not believe them. This delay cost us weeks. How could this happen? To explain our blindness, let's look at the experimental results and compare them to our expectations.

The results from one typical observer are shown in Fig. 2.10 (different panels for different frequencies of background) and again in Fig. 2.11 (results for different frequencies juxtaposed) following the same conventions as Figs. 2.2 and 2.3 for the light-adaptation experiments. Probe threshold (this time  $\Delta C$  rather than  $\Delta I$  however) is plotted versus the phase of the probe and, to help show trends in results, the phases are plotted through two cycles on the horizontal axis.

At low temporal frequencies of flicker, the results were in line with our expectations. In particular, the probe thresholds generally followed the contrast of the background as one might expect from some process that was fast compared to the frequency of the background. This is like the earlier light-adaptation results.

However, look what happened across the range of frequencies. The curves in Figs. 2.10 and 2.11 are on average moving dramatically downward as background frequency increases. They are moving downward rather than staying at roughly the same level (which we had thought they might) or going upward (which we had also thought they might as in the light adaptation results of Figs. 2.2 and 2.3). To put it in other

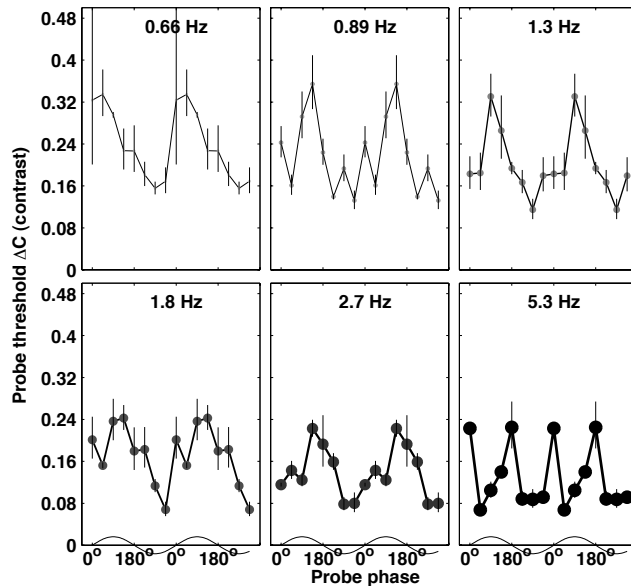


Figure 2.10: Results of the contrast-adaptation experiments for a typical observer (JW). Patterned probe thresholds (vertical axis) are plotted as a function of phase (horizontal axis) for various flicker frequencies of the patterned background (different panels). The data points show the mean  $\pm 1$  SE bars from 3 estimates of each threshold (done in 3 different sessions). In each session trials of all 8 phases were randomly intermixed.

words, as the frequency of the flickering patterned background increases, the observer is generally becoming dramatically more sensitive to the patterned probe. This was completely outside the range of effects we had led ourselves to expect.

#### 2.2.4 Functional fixity, the effects of expectancy, and the Buffy idea

So what did we do? What humans often do. Namely we refused to believe our eyes (refused to believe the new evidence) and believed our preconceptions instead. Psychologists have various names for effects of this sort: mental set, functional fixedness, confirmation bias. In the long course of evolution, one presumes these effects have been useful heuristics. But they clearly can be overdone.

We spent weeks proving once more that there were no artifacts in the experimental procedure, no mistakes in the experimental set-up and no errors in the computer program. We collected results from four observers to make sure that the first was not fooling us (or herself) somehow. We kept getting the same results and kept having trouble believing them. And we certainly could not understand them. And we were sorely tempted to stop this line of experimentation altogether.

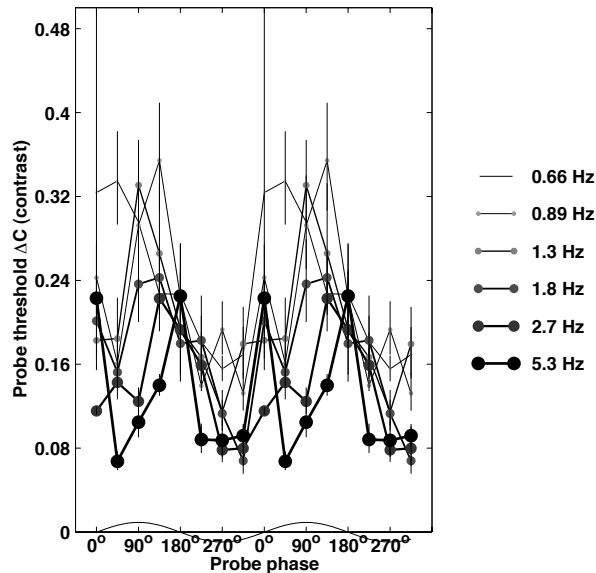


Figure 2.11: The contrast-adaptation results shown in Fig. 2.10 but with the results from all background flicker frequencies superimposed in one panel. In general, the probe thresholds tend to decrease as the frequency of the flickering background is varied from the lowest (lighter-colored and thinner symbols and lines) to the highest (darker and thicker symbols and lines). The light-adaptation results in Fig. 2.3 show the opposite trend with background frequency.

On the other hands, the experimental results were striking, repeatable, and continued nagging at us. And eventually we broke the effects of expectancy. The college-age daughter of one of us (NG) came home for a bad wisdom-teeth extraction operation, and then wanted to spend the rest of Thanksgiving weekend watching *Buffy the Vampire Slayer*. Her mother gave in to daughter's pleas for company. So mother and daughter watched two full seasons of Buffy in four days. Mother sat at the dining-room table occasionally looking at the experimental results of Figs. 2.10 and 2.11 and occasionally sketching diagrams of predictions from the kinds of models outlined above. Daughter lay on the couch occasionally examining the homework she was supposed to be doing. The predictions from mother's preconceived ideas continued to disagree with the experimental results. Buffy continued to slay vampires. But some place during those four days, Buffy turned mother's mind to sufficient mush that the preconceived ideas vanished from it. And a glimmer of an alternate idea took their place. (What happened to daughter's homework is lost to history, but she passed the semester.)

Although this glimmer of an alternate idea took months to become a full-fledged model, it did turn out to account quantitatively for the major features of the experimen-

tal results that were so unexpected to us originally. It now makes the results perfectly believable to us although still somewhat surprising on other grounds. The idea – and its various accompaniments – will be labeled here “Buffy” in honor of its inspiration and for lack of better labels (although in the future we will try to find more traditionally respectable and descriptive names).

It is now hard to believe we were so surprised by these results. Indeed, in retrospect it is hard even for us to understand why. This in our experience is typical also. Once a fixity or expectational set has been broken, people are amazed they were so blind. And forget for the next occasion that they might again be blind.

For the reader to understand the Buffy idea more easily, we think it best to postpone the direct presentation of the idea until later and instead present first some further experimental results we collected to test the idea when it was still a glimmer.

## 2.3 Using Buffy and regular steady-state backgrounds

### 2.3.1 Limits as background frequency gets very low or very high

We constructed two new experimental conditions – called *Regular Steady-State* and *Buffy Steady-State* here – to study what happens at very low and very high frequencies of the flickering spatially-patterned background. To understand these conditions, let's first look at what happens as the flicker frequency gets very low (Fig. 2.12) or very high (Fig. 2.13) while keeping the phase of the probe constant with respect to the flickering background. The contrast of the stimulus seen by the observer is plotted vertically and time is plotted horizontally. Consider an interval of arbitrary length just before the moment at which the probe is added to the background. This interval is shown as a shaded area in these two figures.

Consider Fig. 2.12. As the background flicker frequency gets lower, the function relating the flickering background's contrast to time, which starts out showing several cycles of sinusoidal fluctuation (top of Fig. 2.12), flattens and becomes closer and closer to a horizontal line during the interval just before the probe (bottom Fig. 2.12). Or, to say it another way, as the flicker frequency gets lower, the background stimulus during the before-probe interval becomes more and more like a stimulus of constant contrast. Or, to say it still a third way, the function that relates the flickering background's contrast to time during the interval before the probe reaches a conventional limit: its limit is the function where contrast equals a constant during that interval.

If you look more closely at Fig. 2.12, you can also see that (as background frequency gets lower and lower) the contrast level of the limiting function depends on the phase that the probe was in with respect to the background before the limit was reached. In the example of Fig. 2.12, this phase is approximately midway between the trough (phase  $270^\circ$ ) and the next zero-crossing. We will call the limiting contrast level  $C_{BP}$  (where the  $B$  and  $P$  stand for background and phase); it is shown as a solid horizontal line in the bottom drawing of Fig. 2.12. The dashed horizontal line in each drawing shows  $C_0$ , the mean contrast level of the background.

What happens as the background flicker frequency gets higher (Fig. 2.13), again keeping the probe phase constant? Now the function relating the flickering back-

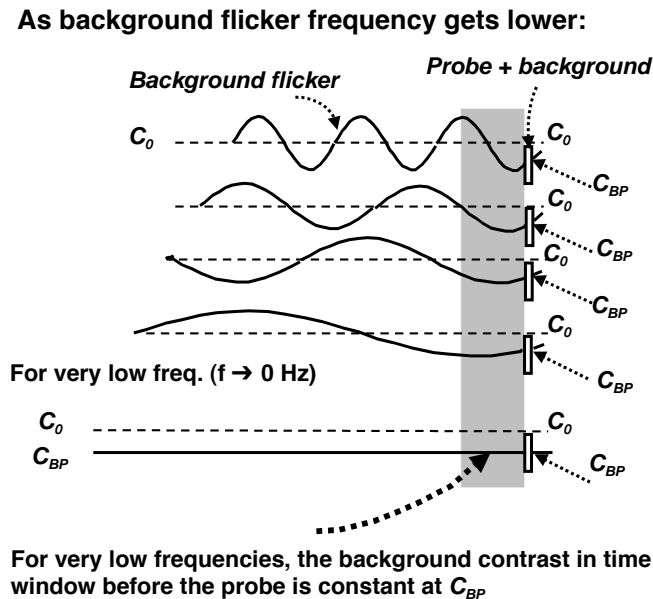


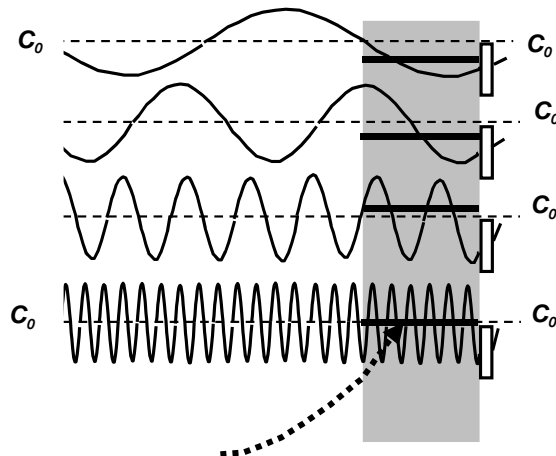
Figure 2.12: A sketch showing the interval before the probe pattern as the frequency of background flicker gets lower and lower while keeping the phase of the probe with respect to the background constant. The little vertical open rectangles represent the probe patterns; the top of the little rectangle is the contrast level of the higher-contrast Gabor patches, and the bottom of the rectangle is that of the lower-contrast Gabor patches. The wavy sinusoidal curve shows the contrast of the flickering background. The dashed horizontal line in each panel shows the mean contrast of the flickering background ( $C_0$ ). The solid horizontal line in the bottom panel shows the contrast ( $C_{BP}$ ) of the flickering background at the phase of the probe. The large shaded rectangular area in this figure simply marks the interval before the probe is presented.

ground's contrast to time does not approach a conventional limit at all. It just oscillates faster and faster (bottom Fig. 2.13).

But what if there is a process that integrates over recent contrast within a temporal window (e.g. the shaded area in Fig. 2.13)? For such a process, the average contrast in the temporal window determines the state of the process at the time the probe is presented. And, to foreshadow later explanations, the existence of such a process forms one important part of the Buffy idea.

In Fig. 2.13, at each frequency of background flicker, the average contrast during the temporal window before the probe is sketched as the thick black horizontal bar extending throughout the shaded area representing the temporal window. As you can see in the figure, as background flicker frequency gets higher and higher (curves from top to bottom), the average contrast in that temporal window goes up and down, but the

**As background flicker frequency gets higher:**



**For very high frequencies, the average contrast in time window before probe settles down at overall mean contrast  $C_0$ .**

Figure 2.13: Sketch showing what happens in the interval before the probe as the frequency of the background flicker gets higher and higher while keeping the phase of the probe with respect to the background constant. Symbols and conventions as in Fig. 2.12 with the following additions. The large shaded vertical rectangle now more specifically indicates the temporal-integration window of a hypothetical process. And the thick horizontal bar in that shaded area indicates the average contrast of the flickering background during that temporal-integration window; thus the thick horizontal bar also indicates the state of the hypothetical process at the moment the probe is presented.

amplitude of the ups and downs gets smaller and smaller. Thus the average contrast in the window does reach a conventional limit. That limit equals the mean background contrast of the sinusoidal flicker  $C_0$  (shown by the dashed horizontal line). And notice that this average contrast settles down at that same contrast *no matter what the phase of the probe relative to the background*.

### 2.3.2 The Regular Steady-State condition – to study very low frequencies

How can we study very low frequencies? The limiting situation shown in Fig. 2.12 motivates our *Regular Steady-State* condition shown in Fig. 2.14. The general Regular Steady-State paradigm is shown in the upper panel of this figure and examples at two

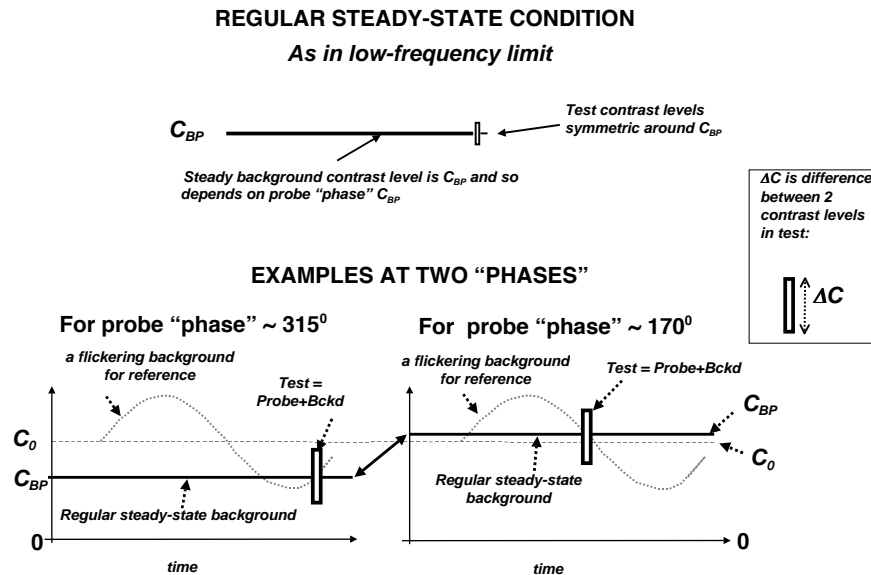


Figure 2.14: Sketch of time course of stimuli in the Regular Steady-State condition. The top panel shows a general case, and the bottom two panels indicate some of the details for two examples “phases”. “Phase” here refers to the phase of the probe with respect to the flickering background (which was kept constant) as the frequency got lower and lower in the limiting process that motivated this steady-state condition (Fig. 2.12).

particular probe “phases<sup>3</sup>” are shown in the lower panels. As you can see in the figure, the Regular Steady-State condition is constructed by using a steady background – at the appropriate contrast  $C_{BP}$  for the “phase” one is considering – instead of using very-low frequency flicker. The test stimulus remains the same as it was on the flickering background.

To summarize: the Regular Steady-State condition has the following two properties:

1. The steady background contrast ( $C_{BP}$ ) changes depending on the “phase” of the probe.
2. For all “phases” the contrasts of the two element types in the test stimulus are symmetric around the steady background contrast.

<sup>3</sup>A terminological aside: The use of the word “phase” in this steady-state condition (and below in another steady-state condition) is a bit odd. Thus we put it in quotation marks for this initial discussion although we will drop the quotation marks in what follows. Note that the “phase” in a steady-state condition is equal to the phase of the probe with respect to the flickering background (which was kept constant) as the frequency changed in the limiting process that motivated the steady-state condition (Fig. 2.12 and, for the next steady-state condition, Fig. 2.13).

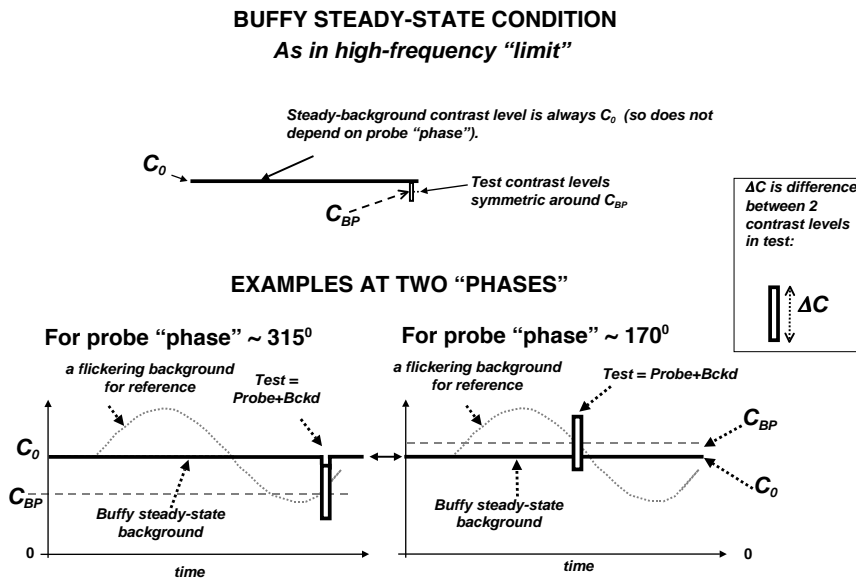


Figure 2.15: Sketch of time course of stimuli in the *Buffy Steady-State* condition. Same format as Fig. 2.14. “Phase” here refers to the phase of the probe with respect to the flickering background (which was kept constant) as the frequency got higher and higher in the limiting process that motivated this steady-state condition (Fig. 2.13).

**Details of the Regular Steady-State condition.** The methods for the steady-state experiments were identical to those for the flickering-background experiments (see above) with the following specifications: The steady-state background was on for about 2.6 seconds and the probe was approximately in the middle. There was always at least 1 second of steady background before and after the probe.

### 2.3.3 The Buffy Steady-State condition – to study very high frequencies assuming a temporal-integration process

How can we study very high frequencies given that it is impossible to generate very high frequency flickering backgrounds? The limiting situation shown in Fig. 2.13 motivates the *Buffy Steady-State condition* shown in Fig. 2.15. A general sketch is at the top and two example phases below.

As you can see in the figure, the Buffy Steady-State condition is constructed by using a steady background instead of the desired very-high frequency flicker. The steady background will be at the same contrast  $C_0$  independent of phase. The test (background-plus-probe) stimulus remains the same as it was on the flickering background.

To summarize, the Buffy Steady-State condition has the following two properties:

1. The steady background contrast is  $C_0$  for all “phases” of probe.
2. The two contrasts in the test are NOT usually symmetric around the steady background contrast (which is always  $C_0$ ). The two contrasts are symmetric around  $C_{BP}$  which changes depending on the phase of the probe.

Note that this condition is only a way to study very high temporal frequencies if there is a process that integrates over a temporal window. And even then it really is only studying certain aspects of behavior at very high frequencies. But, as mentioned above, the temporal-integration process is an important part of the idea we were trying to study.

**Details of the Buffy Steady-State condition.** Other than the differences between the two steady-state conditions portrayed in Figs. 2.14 and 2.15, all methods for the Buffy Steady-State were like those for the Regular Steady-State condition.

### 2.3.4 Experimental results from the two steady-state conditions

The results from the Regular and Buffy Steady-State conditions are shown in the left and right panels, respectively, of Fig. 2.16 for the same observer JW whose results on flickering backgrounds were shown earlier (in Figs. 2.10 and 2.11). Fig. 2.17 shows the results from Fig. 2.16 again; now they are superimposed on this observer’s results for the flickering backgrounds shown earlier.

As can be seen in the figures, the results are rather as you would expect by extrapolation from the results on the flickering backgrounds to even lower and even higher frequencies. In particular, the Regular Steady-State thresholds are quite high like those on the lowest-frequency flickering backgrounds; the Buffy Steady-State thresholds are very low, somewhat lower even than the highest flicker frequency studied for this observer JW (5.3 Hz).

Another aspect of the Buffy Steady-State results is worth pointing out. They, like those on flickering backgrounds at the highest flicker frequencies, show almost complete frequency doubling. (The probe-threshold curves show two almost-identical maxima and two almost-identical minima per cycle of the background flicker.) Such complete frequency doubling (visible in every individual session from every observer) was much more dramatic than any frequency-doubling seen in the light-adaptation results and had initially surprised us when we had only collected the flickering-background results and had not yet had the glimmer of the Buffy idea.

### 2.3.5 Examining Buffy Steady-State results further: the extreme phases

Consider the phases that produce the extreme thresholds in the experimental results from the Buffy Steady-State condition. Phases at or near the zero-crossings ( $0^\circ$  and  $180^\circ$ ) lead to the maximal thresholds, and phases at or near the peak ( $90^\circ$ ) and trough ( $270^\circ$ ) lead to minimal probe thresholds (Fig. 2.16). (Remember that “phase” here

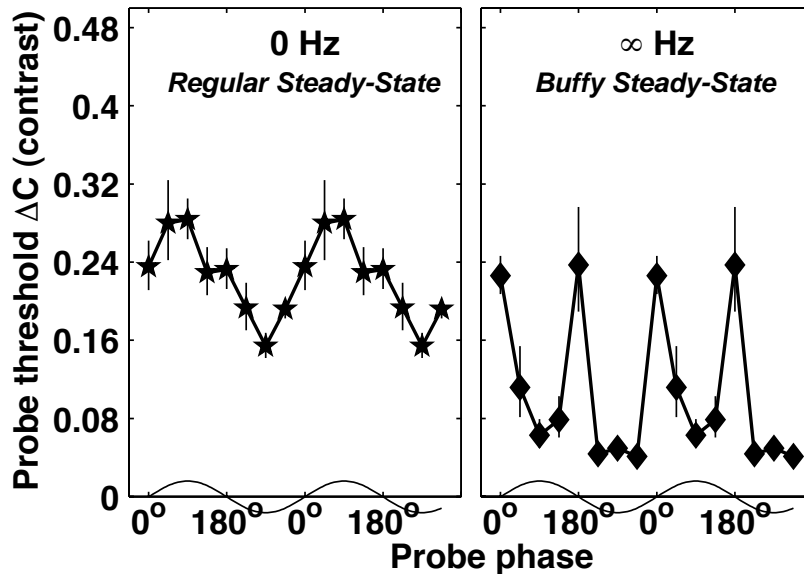


Figure 2.16: Results from the Regular (left panel) and Buffy (right panel) Steady-State conditions for the observer JW whose results on flickering backgrounds are shown in Figs. 2.10 and 2.11. The results shown are the mean  $\pm 1$  SE bars from 6 estimates of each threshold (done in 6 different sessions). In each session all eight phases were studied in randomly intermixed trials.

refers to the phase of the probe with respect to the background flicker in the limiting process shown in Fig. 2.13.)

The stimuli corresponding to these extreme phases are illustrated further in Fig. 2.18. The background stimulus (a steady field of contrast  $C_0$ ) is presented to the observer for some period of time. This stimulus could also be called an “adapting stimulus” and is so called in Fig. 2.18. In the experiments here, the duration of the adapting stimulus happened to be a bit more than 1 second, but anything that long or longer probably produces much the same result (and things somewhat shorter might also). Then the test stimulus (the combination of background-plus-probe) is presented briefly.

For the peak phase ( $90^\circ$ ), the average of the two contrast levels in the test stimulus is ABOVE the background contrast. Indeed usually it turns out that not only the average but both of the individual contrast levels in the test stimulus at threshold are above the background contrast.

For the zero-crossing phases ( $0^\circ$  and  $180^\circ$ ), the average contrast of the test stimulus is equal to that of the background stimulus, and the two levels at threshold therefore STRADDLE the background contrast.

For the trough phase ( $270^\circ$ ), the average of the two contrast levels in the test stim-

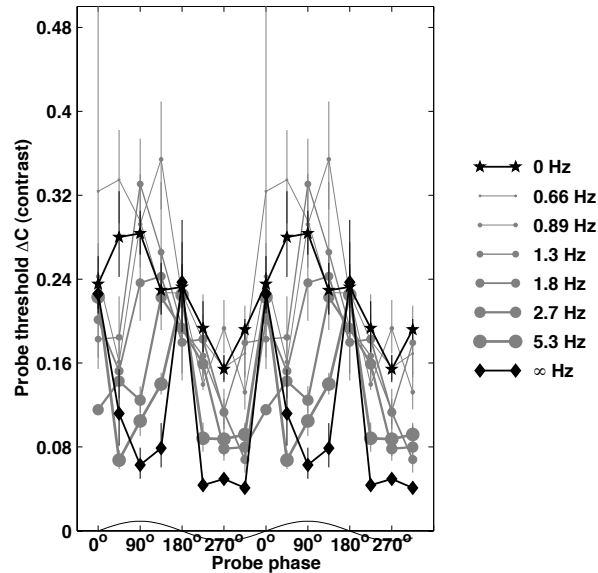


Figure 2.17: The steady-state and flickering-background contrast-adaptation results replotted from Figs. 2.11 and 2.16.

ulus is BELOW the background contrast. Usually it turns out that not only the average but both of the individual contrast levels in the test stimulus at threshold are below the background contrast.

Fig. 2.19 displays numerically some results from the Buffy Steady-State for these three extreme phases for two observers (JW, the observer whose results have been shown in earlier figures, and another observer SH). The contrast levels in the test stimuli at threshold are given in one column and then the difference between those contrasts is given in the adjacent column.

As one would expect from the curves in Fig. 2.16 right panel, the contrast difference at threshold is much larger for the test stimulus which STRADDLES the background (zero-crossing phase) than for either of the other test stimuli ABOVE or BELOW the background contrast. And the threshold contrasts for those other two are very similar to one another but on opposite sides of the background contrast. The observers cannot see that straddling stimulus well enough to identify the orientation of its stripes until the two contrast levels defining the stripes are very far apart, e.g. about 23% for observer JW and 41% for observer SH. For the test stimuli where both contrast levels are on one side of the steady background contrast, the threshold contrast differences range from 5% to 8%.

The results are particularly dramatic for observer SH. For this observer, the two contrast levels (70% and 30%) in the threshold for the STRADDLING test stimulus

### Three Phases in the Buffy Steady-State condition

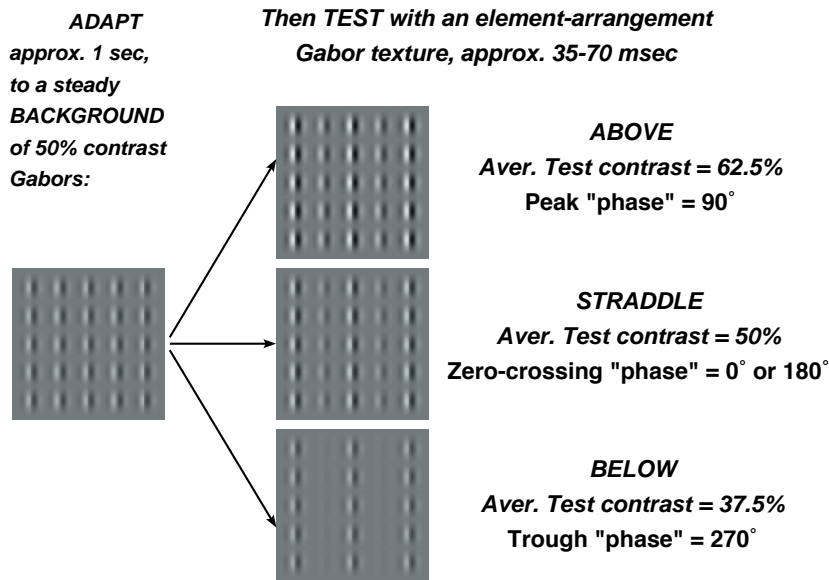


Figure 2.18: Illustrations of three kinds of trials (the three extreme phases) in the Buffy Steady-State condition. The observer sees ("adapts") to a steady background of equal-contrast Gabor patches for a little more than 1 second and then sees a briefly-presented test stimulus which has either vertical contrast-defined stripes (as in the figure) or horizontal. The observer is to say which orientation the stripes are. Notice that the test stimuli have patches of two contrast levels. Depending on the phase, these contrast levels can (1) average ABOVE, or (2) STRADDLE, or (3) average BELOW the contrast  $C_0$  of the Gabor patches in the steady background.

are literally further out from the background contrast level than are any of the contrast levels in threshold test stimuli ABOVE (66%, 59%) or BELOW (41%, 34%) the steady background.

To look at the results in a slightly different way: Which test stimulus is hardest for the observer? It is the middle one – the test stimulus in which the contrasts are STRADDLING the background contrast. Thus, when looked at in this way, the contrasts NEAREST to the adapting contrast are actually hardest for the observer to process (in whatever way is necessary for the observer to do this identification task). This seemed a surprising effect of adaptation to us, and we will return to this topic again in the last section. But first we will present a model that can explain the mechanism of this adaptation effect.

		<b>Observer JW</b>		<b>Observer SH</b>	
<b>TEST</b>		<b>Test contrasts at threshold</b>		<b>Test contrasts at threshold</b>	
<b>ADAPT</b>	<b>ABOVE</b> 90°	c1: 66% c2: 59%	diff = 6%	c1: 66% c2: 59%	diff = 8%
	<b>STRADDLE</b> 0°, 180°	c1: 62% c2: 38%	diff = 23%	c1: 70% c2: 30%	diff = 41%
	<b>BELOW</b> 270°	c1: 40% c2: 35%	diff = 5%	c1: 41% c2: 34%	diff = 7%

Figure 2.19: Results from two observers for the three extreme phases of the Buffy Steady-State condition illustrated in Fig. 2.18. For each observer, the contrast levels in the test stimuli at threshold are given in one column and then the difference between those contrasts in the adjacent column. (All contrasts were rounds to the nearest 1%. Since the thresholds plotted at 0° and 180° in Fig. 2.16 are actually from the same straddling test stimulus, these were averaged together for the numbers displayed for the straddling test stimulus here.) Notice that the threshold for the stimulus that STRADDLES the adapting level is much higher than that for stimuli either ABOVE or BELOW the adapting level.

## 2.4 A new kind of complex channel with embedded contrast adaptation (the Buffy channel)

Let's return now to the glimmer of an idea that appeared during the days of watching *Buffy the Vampire Slayer*, the idea that helped us begin to understand the flickering-background results that we had found so surprising (e.g. Figs. 2.10 and 2.11). Fig. 2.20 attempts to portray the idea as it first appeared. It was a vague idea that some adaptable "contrast comparison process" might be operating in these perceptions, a comparison process with the following characteristics:

The input to the possible comparison process is some kind of response that measures the local contrast at different positions in the visual field. For the texture patterns used here, measuring the local contrast is like measuring the contrast in each Gabor-patch element.

There is a "comparison level" (at each position in the visual field) and the process compares the current local contrast to that comparison level. The comparison level is shown by the arrow in the figure.

## A possible comparison process

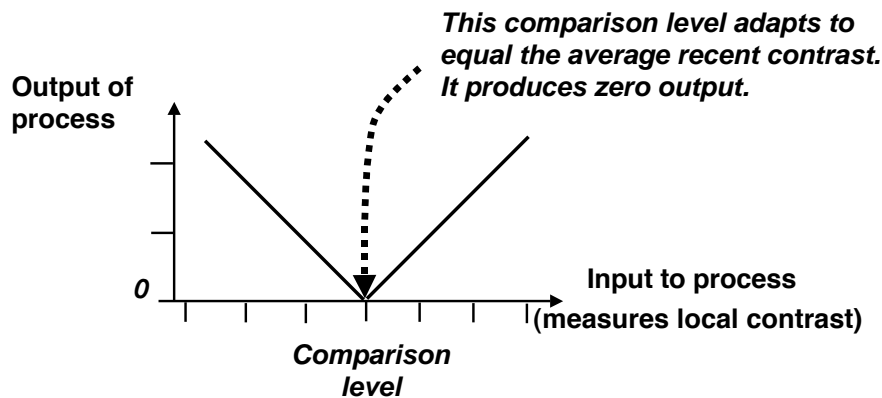


Figure 2.20: Diagram of a comparison process – an idea leading to a possible explanation for results like those in Fig. 2.19.

And, most importantly, the comparison level is adaptable: It adapts to equal the average of its recent input (the input in some temporal interval before the current moment). Or, in other words, the comparison level equals the average recent local contrast.

The output of this possible comparison process (at any particular position) is then a measure of the difference between the current contrast and the average recent contrast (at that position). The output is zero when the current contrast and average recent contrast are equal, and it is positive when they are different, where the magnitude depends on just how different they are, but the direction of the difference is ignored. To get from this notion sketched in Fig. 2.20 to something rigorous required some further steps.

### 2.4.1 How to measure local contrast for the input to this process

First, where could the input to this comparison process come from? That is, how can we measure local contrast to use as the input to this process? Fig. 2.21 shows a structure that has been used by a number of people for this purpose. This particular filter-nonlinearity-filter (*FNF*) structure computes a measurement of the contrast in a particular spatial-frequency and orientation range (that corresponding to the first filter's receptive fields) at each position in the visual field.

The second filter in this structure is simply an excitatory-center-only receptive field. Thus it blurs the rectified output from the first filter. The second filter's excitatory-only center is about the size of a receptive field at the first filter, and thus it blurs over an area about the size of one such receptive field – which is approximately the size of the Gabor patches in our stimuli.

## TO MEASURE LOCAL CONTRAST (e.g. contrast in Gabor patch)

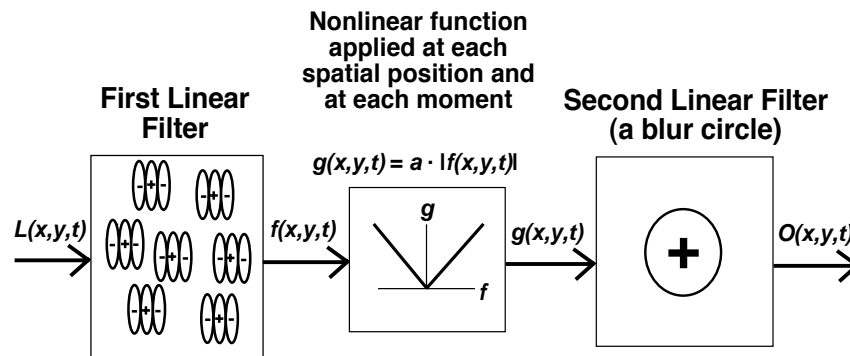


Figure 2.21: Diagram of a structure that can measure local contrast at the spatial frequency and orientation range corresponding to the receptive fields of the first-stage filter. For the patterns used in the contrast adaptation experiments here (represented by the function  $L(x, y, t)$  in the figure), the output  $O(x, y, t)$  of this structure (at any particular spatial position) will be a good measure of the contrast of a Gabor patch at that position (assuming the Gabor patch has spatial frequency and orientation appropriate for the first-stage filter). Functions specifying the outputs at intermediate stages within the channel are represented by the symbols  $f$  and  $g$ , and  $a$  is an arbitrary constant. See text for further description.

You would need multiple  $FNF$  structures like the one in Fig. 2.21 to measure the local contrast at all ranges of spatial frequency and orientation.

### 2.4.2 Extending the original complex channel to include this comparison process

Next we need to add an adaptable contrast comparison process (like that in Fig. 2.20) to the original complex channel (Fig. 2.9). The resulting Buffy channel (as we will call it for this chapter) is sketched in Fig. 2.22.

A Buffy channel has three linear filters with two nonlinearities sandwiched between (a  $FNFNF$  structure). The first three stages ( $F1 - N1 - F2$ ) have a structure identical to that in Fig. 2.21. The output from  $F2$  – called  $O_1(x, y, t)$  in Fig. 2.22 – is a measure of local contrast for a particular range of spatial frequency and orientation. It is fed into a second nonlinearity  $N2$  that acts at every spatial position. Very importantly, this

## A Buffy Channel

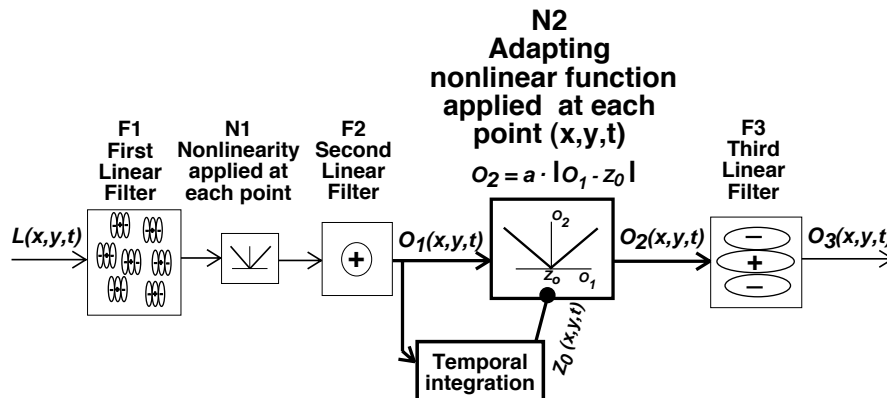


Figure 2.22: Diagram of a new kind of complex channel – a Buffy channel. It contains an embedded contrast adaptation process. The first three processes in the Buffy channel are a repeat of the structure shown in Fig. 2.21. Their output, labeled  $O_1(x, y, t)$  or  $O_1$ , is a measure of the local contrast at position  $(x, y)$  and time  $t$ . The zero-point (or comparison level) of the second nonlinearity  $N2$  (the adapting nonlinearity) is denoted by the symbol  $Z_0(x, y, t)$  or  $Z_0$ . This zero-point is set equal to the output of the temporal integration box, that is, to the recent time-averaged output of the second filter  $F2$ . The zero-point thus reflects the recent time-averaged local contrast. Thus  $O_2(x, y, t)$  or  $O_2$  (the output of the second nonlinearity  $N2$ ) is zero whenever the current stimulus contrast is the same as the recent time-averaged contrast. And the output becomes larger as the current contrast becomes further (either increasing or decreasing) from the time-average contrast. The nonlinear function in  $N2$  is shown here as a full-wave rectification (with a constant of proportionality  $a$ ), but see Fig. 2.23 for more information. The output  $O_3(x, y, t)$  of the third filter  $F3$ , which is also the output from the whole Buffy channel, depends on the arrangement of the two element types in the pattern. The third filter  $F3$  shown in the illustrated Buffy channel will respond to contrast-defined horizontal stripes at a spacing corresponding to the spacing between the excitatory and inhibitory parts of the receptive field shown for  $F3$ . See text for further description.

second pointwise nonlinearity  $N2$  is *not* stable across time but instead adapts. In fact, it will play the role of the comparison process that was sketched in Fig. 2.20. So let's look at this adaptable nonlinearity  $N2$  more carefully.

The form of the nonlinear function  $N2$  at any point  $(x, y, t)$  is some type of rectifi-

ation. (An approximately symmetric full-wave rectification is shown in Fig. 2.22.) Note the temporal-integration box in the figure: it integrates the measure of local contrast  $O_1(x, y, t)$  where the integration is primarily over time although there could be some spatial integration as well. Thus the output of the temporal integration box  $O_2(x, y, t)$  represents the average contrast over the recent past at position  $(x, y)$ . This output is then connected (by a funny round-headed “arrow”) to the second nonlinearity  $N2$ , in particular, to its zero-point  $Z_o$  which is the level of input into  $N2$  that leads to an output of zero. The round-headed “arrow” is meant to indicate that the output from the temporal-integration box directly sets the value of the zero-point, and thus the zero-point  $Z_o$  at each position  $(x, y)$  acts like the comparison level in Fig. 2.20. It adapts to reflect the average recent contrast at that position. Therefore, if the current contrast at a particular position equals the recent averaged contrast at that position, the output of the second nonlinearity  $N2$  will be (approximately) zero there. But if the current contrast differs from the recent averaged contrast (either above or below it), the output from  $N2$  will be positive with a magnitude reflecting the difference between the current contrast and the recent averaged contrast.

The output from the second nonlinearity  $N2$  is fed into a third filter  $F3$ . This third filter in the Buffy channel performs the function of the second filter in the original form of complex channel (Fig. 2.9). It is able to respond to the striped arrangement in an element-arrangement texture when the stripes have the right spatial period and orientation to match the characteristics of its receptive field and the elements are matched to the characteristics of the first-stage receptive fields.

Fig. 2.22 shows one Buffy channel. Of course, we are assuming the existence of many Buffy channels sensitive to many different ranges of spatial frequency and orientation at the first and third stage filters. The excitatory center in the second filter is always approximately matched in size to the receptive field of the first filter.

As it turns out, we will also have to assume that the kind of rectification in the adapting nonlinearity  $N2$  can vary depending on the observer. It is usually not a standard full-wave or half-wave rectification but something intermediate.

And, further, for a half-wave or intermediate case, we will need to assume a pair of Buffy channels, identical to one another except for the exact form of their second nonlinearities. Fig. 2.23 shows an example of the  $N2$  functions for such a pair of channels. In each such pair, one channel’s second nonlinearity has an asymmetry favoring increases or onsets in contrast relative to the comparison level (called here the “On” member of the pair); the other’s second nonlinearity has an asymmetry favoring decreases or offsets in contrast (the “Off” member of the pair). We are using “On” and “Off” here to refer to contrast in analogy to the way they have often been used to refer to increases (onsets) and decreases (offsets) of luminance.

Before moving on to predictions for experimental results, let’s note one general aspect of the change in the second nonlinearity  $N2$ , that is, the resetting of the zero-point (i.e. comparison level) based on recent average contrast. This change is of an additive/subtractive sort rather than multiplicative/divisive. The zero-point that is changed is *subtracted* from the current contrast. This change is not in a multiplicative or divisive “gain” parameter, which is the kind of change in most contrast-gain processes that people talk about (including the normalization process that was the beginning point of our interest in these experiments).

## N2 functions from a pair of otherwise-identical Buffy channels

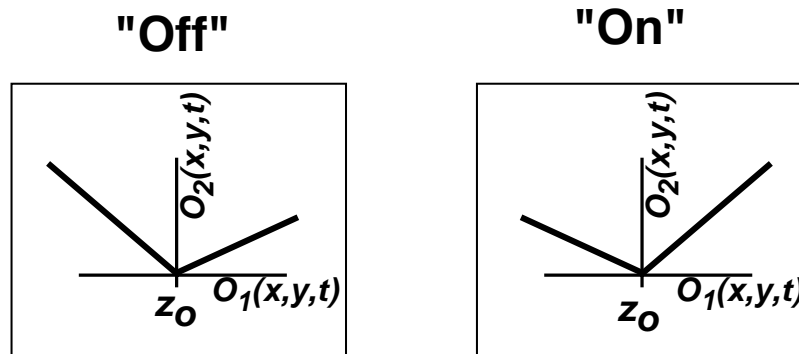


Figure 2.23: A possible pair of  $N2$  functions for a pair of Buffy channels. For most observers, the static form of  $N2$  does not seem to be either precisely a full-wave or half-wave rectification but rather something intermediate. And, further, one needs to assume that there will be a pair of channels where one channel has an asymmetry favoring increases in contrast (the "On" member of the pair) and the other has an asymmetry favoring decreases in contrast (the "Off" member of the pair). The  $N2$  functions from such a pair are what is shown in this figure.

The next figure (Fig. 2.24) shows the full model framework that appeared earlier in Fig. 2.8 but now with the original complex channels replaced by Buffy channels.

### 2.4.3 Predictions from Buffy channels

Fig. 2.25 compares experimental results (filled diamonds) with predictions from Buffy channels (thick gray smooth curves). The results shown are from the Buffy Steady-State conditions for four observers (including JW and SH from earlier figures). The predictions are from a model containing Buffy channels and little else except the necessary rule relating the observer's response to the outputs of the channels. (We will get back to other processes in Fig. 2.24 – e.g. normalization – below.) And we made one important further assumption for the predictions shown in Fig. 2.25: We assumed that the integration time of the temporal-integration box was shorter than the approximate 1 second of steady-state background before the probe was presented. Thus, the zero-point of  $N2$  for these predictions always corresponded to the Buffy Steady-State

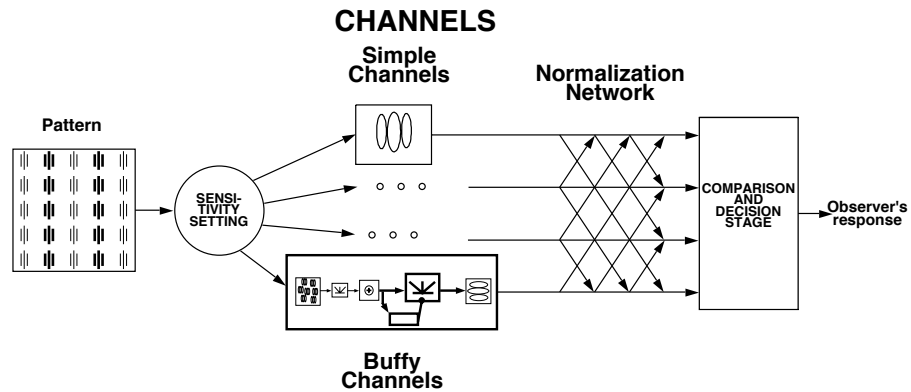


Figure 2.24: The full model from Fig. 2.8 with Buffy channels replacing the original complex channels.

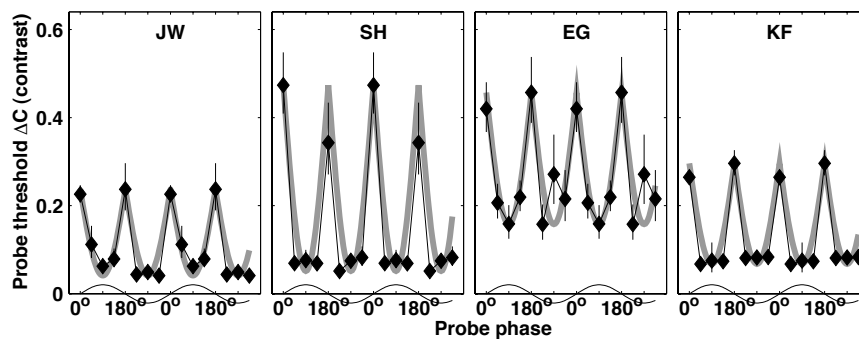


Figure 2.25: The solid diamonds show results from the Buffy Steady-State experiments for 4 observers (in different panels). The thick gray curved lines are predictions from a simplified model containing only Buffy channels. The predictions describe the experimental results well.

background contrast level  $C_0$  (which was a contrast of 0.5 or 50% in these experiments).

The predictions in Fig. 2.25 can be described quite easily as follows (although we will not derive them here). You take a sinusoidal function of the same frequency as the background flicker and in phase with it. You full-wave rectify this sinusoidal function. You turn it upside down. You move it up and down (an additive parameter) and stretch it or compress it uniformly (a multiplicative parameter). There is another possible degree of freedom in the predicted functions, although we didn't use it here: It allows you to

cut off the rounded bottom of the predictions with a flat horizontal line at an arbitrary height. In principle, you do all of these things simultaneously until you get a best fit. Here we have just settled for a satisfactory fit by eye since that fit is already better than the experimental results collected to date could reject.

For the interested reader, the next subsection will pursue somewhat further the effect of changing various characteristics of the Buffy channels on the exact predictions. But a reader can skip this next subsection with no loss of continuity.

#### 2.4.4 The parameters of predictions like those in Fig. 2.25

Collectively the freedom in the manipulations described above for fitting the predictions to the results in Fig. 2.25 come from 3 parameters in the simplified Buffy channel model. These are (1) a parameter representing the observer's visual sensitivity to these patterns, (2) a parameter representing properties of the decision stage, and (3) a parameter indicating the exact form of the functions characterizing  $N2$ .

This last parameter is most interesting. If these  $N2$  functions are exactly conventional full-wave functions, the maximum thresholds (those for the STRADDLING test stimuli – phases  $0^\circ$  and  $180^\circ$ ) would be infinite. This is unlike any of our observers so far. If the  $N2$  functions are exactly conventional half-wave functions, the ratio between the maximum thresholds (for STRADDLING test stimuli) and the minimum thresholds (for ABOVE and BELOW test stimuli at phases  $90^\circ$  and  $270^\circ$ ) would be at most equal to 2. The exact ratio depends on factors other than the functions in  $N2$ . Conventional half-wave functions may predict the results for observer EG for whom the ratio of maximum to minimum is the smallest of all the observers (and not significantly different from 2). For the other observers we have studied so far, however, the  $N2$  functions need to be intermediate between half-wave and full-wave (like those in Fig. 2.23). Alternately, although we have not yet explored this formally, the results might also be predictable by mixtures of different kinds of channels, where some channels have half-wave and some have full-wave  $N2$  functions.

Still other degrees of freedom in the model are connected with other variations in the form of the  $N2$  functions. For the predictions in Fig. 2.25, we considered only piecewise-linear functions and assumed that the two functions in a pair are mirror images of each other (as in Fig. 2.23). However, some previous evidence (Graham and Sutter, 1998) suggests that either these  $N2$  functions – or else the functions in  $N1$  – need to be power functions with an exponent between 3 and 4 rather than piecewise-linear (an exponent of 1) as shown here. If the  $N2$  functions are such power functions, the predictions in Fig. 2.25 would be somewhat differently shaped but they would still fit the results well. Or to put it another way, the results shown here cannot discriminate among different powers (unlike the results in Graham and Sutter, 1998).

There is another variation in the exact form of the  $N2$  functions which results like those in Fig. 2.25 could discriminate among. The pairs of functions we have considered so far have contained mirror-image members (as in Fig. 2.23). If the two members of a pair are not mirror images, the thresholds for the peak ( $90^\circ$ ) and trough ( $270^\circ$ ) phases may differ from one another. There is a hint of such asymmetry in the data of JW (Fig. 2.25) but that result is not reliable, and we have not seen any observers showing a reliable difference.

## 2.5 Where are we now?

Overall, the fits in Fig. 2.25 – which assume the contrast-adaptation process embedded in Buffy channels – are very good. This successful prediction of experimental results suggests that such an adaptation process exists. But that leaves us with a number of questions.

### 2.5.1 Do we know anything yet about the dynamics of the contrast-adaptation process in the Buffy channels?

First, let's look at the following question. Can we use the experimental results on flickering backgrounds (Fig. 2.10) to give us some indication of the dynamics of the adaptation process in the Buffy channel (the process that resets the comparison level  $Z_o$  in the adapting nonlinearity  $N2$ )?

Remember that the results from the Buffy Steady-State condition are presumed to be the results on an infinitely high frequency flickering background (assuming the existence of a temporal-integration process). Notice also the results on a 5.3 Hz flickering background (lower right panel Fig. 2.10) show strong indications of the same process that is shown in the observer's results in the Buffy Steady-State condition (Fig. 2.16 right panel); in particular, there is dramatic frequency doubling in both. This similarity of results suggests that 5.3 Hz is approximately equal to an infinitely-high frequency (only for our purposes here, of course). Or to put it another way, first noting that one cycle of 5.3 Hz occupies 190 msec, the similarity suggests that most (although perhaps not all) of the temporal integration in the Buffy channels extends for at least 190 msec.

Can we estimate an upper limit on how long the temporal integration window might be? For some observers (but not JW), there is also some obvious frequency doubling in the results for a frequency of 2.7 Hz or even lower. And even when you cannot see frequency doubling in the results, there may be some effect of the contrast-adaptation process in the Buffy channels nonetheless. (The frequency-doubling effect of a Buffy channel's adaptation process may be counteracted by effects of some other process controlled by contrast, for example the normalization which was ignored in the predictions of Fig. 2.25. We will return to that normalization process below.) However, we think it is probably reasonable to say the following, first noting that one cycle of 2.7 Hz is 380 msec. Although there may be some effect of temporal integration in the Buffy channels that extends for as long 380 msec or even longer, most of the temporal integration is over before then. So the effective integration time of the adapting nonlinearity in the Buffy channels might be something like 250 msec.

It is interesting to consider the relationship of this timing to that of eye movements. The integration time here is probably short enough that the comparison level (zero-point) in the Buffy channel is totally reset by the end of most fixations of the eye. And resetting a contrast comparison level within each eye fixation may be important since the contrast (in the image stimulating any one receptive field) is almost completely uncorrelated from one fixation to the next (Frazor and Geisler, 2006).

But more work – both empirical and modeling – is certainly required to tie this estimate down. Some of the complications are discussed next.

### 2.5.2 What about the dynamics of the normalization process?

As mentioned before, the predictions in Fig. 2.25 were done assuming the existence of Buffy channels but ignoring the possibility of any kind of process like normalization. Yet we have good reason to suppose that a normalization-type contrast-gain control process is operating on these patterns. In the interest of closure, although more empirical and modeling work needs to be done before we have any good basis for believing anything, let us say what is our current opinion is. We think that some of the changes in probe threshold as we change the phase of the probe and the frequency of the flickering background (especially at very low frequencies) may reflect the dynamics of the normalization process (the process that we intended to study when we began these experiments). Whether it will ever be possible to disentangle the dynamics of normalization from those of the adaptation process in the Buffy channel is unclear to us.

### 2.5.3 What about the original form of complex channels?

When one finds a new result, which requires for its explanation the addition or modification of a process in a model that has already successfully predicted many other experimental results, one needs to go back and think carefully about the previous successful predictions.

One very important issue is whether the new entity – in this case the Buffy channel – is still consistent with successful predictions for the older experiments. In our case, the answer is quite easy. Since all our older experiments had the observer adapted to zero contrast (a plain gray field) before each test stimulus, that is the only condition we need to consider. The Buffy channels, after adaptation to zero contrast, act exactly like the original form of complex channel. Since the original form led to successful predictions, so will the Buffy form.

A more subtle question is whether the human visual system might contain both the original *un*-adaptable type of complex channel and also the adaptable Buffy channels. For reasons that are not worth the space to explain here, we think the answer is yes. We think we can design experiments that will test the possibility that both exist, but have not done so yet.

### 2.5.4 Why should there be a contrast-adaptation process like that in the Buffy channels?

Let's go now to the evolutionary "why" question. Why does a process exist that produces results like those in Figs. 2.17 and 2.19, results that can be predicted by the Buffy channels shown in Fig. 2.22? Why should there be a process like this at all? Why has evolution led to this? Getting definitive evidence about any evolutionary "why" question is difficult if not impossible. But thinking about the functionality – about the evolutionary history – of visual processes has led to many interesting insights and ideas for further research. So we will do a bit here.

Adaptation to many different characteristics of visual stimuli occurs. We have discussed adaptation both to luminance and to contrast in this chapter, but there are a very

large number of others. A number of different reasons why such adaptation might exist (functions that such adaptation might perform) have been suggested, most of which can be grouped within two general classes. First let's briefly look at these two classes and then go on to ask whether either class of explanation can help make sense of the experimental results here.

1. *Re-centering the operating range of a process to enhance its ability to discriminate among stimuli within that operating range.*

A system can usually only respond in a well-differentiated way to stimulus values within an operating (or dynamic) range, that is, within a limited portion of the possible range of stimulus values in the environment. The responses to values below and above that operating range are all at the minimum or all at the maximum response level, respectively. Thus the observer can discriminate very well between two values that are within the operating range, but cannot discriminate at all well between two values that are either both below or both above that range. But if the placement of that range on the stimulus dimension can be adjusted by a process of adaptation to match the environment, then this limit in dynamic range may not be a large deficit because the range might be approximately in the right place for whatever stimulus comes next. Thus one common idea is that: the function of adaptation is to re-center the operating range of the system to be centered near or at the current adaptation level (the average level in the recent past of whatever kind of input is at issue). The function of light adaptation is widely believed to be of this sort.

2. *Responding to changes in the visual field (because changes signal important events in the environment, and/or to make neural coding more efficient).*

An alternate view is that the important function of adaptation is to suppress the response to unchanged visual stimuli and thereby highlight the responses to changes in stimuli. Two different classes of supporting arguments seem to be given. One is that a change in the visual stimulus is likely to mean an important change in the environment. The arrival of a predator is a common and dramatic example. Alternatively, people sometimes argue that physiology places serious constraints on how much information can be transmitted by the visual system. They argue that a coding scheme that suppresses the amount of information transmitted about unchanged things and only transmits information about changes can be more efficient in various ways. Either of these two classes of supporting argument suggest that evolutionary pressures might tend to produce visual processes that respond best to changes in stimulation rather than to continuation of unchanged stimulation. Or, as sometimes said, the visual system responds best to transients.

Let's look now at whether either of these classes of proposed explanations of adaptation's function helps us make sense of our results here.

The first class of explanation – re-centering the dynamic range at the recent average level in order to keep discriminability high there – does not seem to help us at all. In particular, as mentioned earlier, performance in the Buffy Steady-State condition is

worse for a test stimulus that STRADDLES the background level (the recent average level) than for test stimuli containing contrasts further away from the background (the ABOVE and BELOW test stimuli). Being worst near the adaptation level is exactly the opposite of what you would expect from the first class of explanation.

The second class of explanation may provide some insight although it still leaves us wondering. The adaptation process we are suggesting resets the comparison level in the Buffy channel to represent the recent average level of contrast. Or, to say this in slightly different words: the zero-point of the Buffy channel's second nonlinearity  $N_2$  (at any particular spatial position) resets during the course of a steady background so that the second nonlinearity  $N_2$  produces zero response to a continuation of the steady background. But when a test stimulus comes on, producing a sudden transient in the contrast (at that spatial position), the second nonlinearity  $N_2$  (at that spatial position) now can easily signal that transient change by responding above zero. So the desirability of responding only to changes (and not responding at all to ongoing stimuli), as assumed by the second class of explanation, does seem illuminating up to this point. The adaptation of the zero-point  $Z_o$  in the second nonlinearity  $N_2$  could be a process whose function is to signal transients.

A puzzle remains, however. The experimental results show that the STRADDLING test pattern is very hard to see. In other words, they show that increases and decreases in contrast do not produce different enough results to allow the observer to perceive a pattern made up of these increases and decreases (e.g. our STRADDLING test pattern). To put it still another way, although the fact of a change of contrast can be well signaled in each local area of the pattern, the direction of that change is not signaled well. In terms of our model, increases and decreases in contrast lead to such similar responses by the second nonlinearity  $N_2$  that the contrast-defined arrangement in the test pattern is very hard to perceive.

This last aspect of the Buffy channel and its embedded contrast adaptation seems odd. Why would evolution have led to this result? Was it something that had some particular functional value in and of itself that led to it being actively selected-for by evolutionary pressures? We cannot imagine what that functional value would be, but perhaps someone else can. Or was its evolution a side-effect of the selection of some other thing? That seems a bit more likely to us. The thing that might have been so useful that it drove evolution might well have been the fact of signaling a change, any change, as quickly as possible. And then the following seems plausible to us. Perhaps wiring a neural system so that it can signal a change quickly without regard to direction of the change is much less costly (in terms of whatever kinds of costs limit evolution of neural tissue) than wiring a system to signal quickly both a change and its direction. We wish we had some idea of what the constraint producing that cost might be.

### 2.5.5 Ending comment

And so ends the first episode of a story that began with some experimental results that we did not believe and that might have halted our investigation into the dynamics of contrast-controlled adaptation processes. But we were saved by the expectancy-overcoming effects of watching *Buffy the Vampire Slayer*. This story ends, for now, with the introduction of Buffy channels, a suggested new form of channel that contains

a nonlinearity showing fast contrast adaptation. This contrast adaptation changes an additive constant, a constant you might call “a comparison level”. The integration period for this adaptation may be something like 250 msec. We still find this proposed channel somewhat puzzling. In the near future we plan to investigate it further, empirically and theoretically. We will perhaps find a more descriptive name for it once we understand better its characteristics and function, but for the moment we are happy to give Buffy the credit she deserves.

## Acknowledgements

This work was supported in part by National Eye Institute grant EY08459. Some of these results were presented at the Spring 2005 VSS meeting (Wolfson and Graham, 2005b). We thank our observers for their hours of effort, and we thank Alisa Surkis and Jiatao Wang for computing predictions from the models of contrast-gain processes mentioned in Section 2.2. Finally, we would also like to thank the colleagues who spent time and energy talking to us about these results and ideas.

## References

- Beck, J., Prazdny K. & Rosenfeld, A. (1983). A theory of textural segmentation. In J. Beck, B. Hope & A. Rosenfeld (Eds) *Human and Machine Vision*. pp. 1-38. Academic: New York.
- Carandini, M., Heeger, D. J. & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.*, 17: 8621–8644.
- Frazor, R. A. & Geisler, W. S. (2006). Local Luminance and Contrast in Natural Images. *Vis. Res.*, 46: 1585–1598.
- Graham, N. (1989). *Visual Pattern Analyzers*. Oxford University Press: New York.
- Graham, N. (1992). Breaking the visual stimulus into parts. *Cur. Direct. Psycholog. Sci.*, 1: 55–61.
- Graham, N., Beck, J. & Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation. *Vis. Res.*, 32: 719–743.
- Graham, N. & Hood, D. C. (1992). Modeling the dynamics of light adaptation: The merging of two traditions. *Vis. Res.*, 32: 1373–1393.
- Graham, N. & Sutter, A. (1998). Spatial summation in simple (Fourier) and complex (non-Fourier) texture channels. *Vis. Res.*, 38: 231–257.
- Graham, N. & Sutter, A. (2000). Normalization: Contrast-gain control in simple (Fourier) and complex (non-Fourier) pathways of pattern vision. *Vis. Res.*, 40: 2737–2761.
- Graham, N. & Wolfson, S. S. (2001). A note about preferred orientations at the first and second stages of complex (second-order) texture channels. *J. Opt. Soc. Am. A*, 18: 2273–2281.

- Graham, N., Wolfson, S. S. & Chowdhury, J. (2001). A comparison of light adaptation results from 40 years of the probed-sinewave paradigm. *Invest. Ophth. Vis. Sci.*, 42: S157, abstract #840.
- Hayhoe, M. M., Levin, M. E. & Koshel, R. J. (1992). Subtractive processes in light adaptation. *Vis. Res.*, 32: 323–333.
- Hood, D. C. & Graham, N. (1998). Threshold fluctuations on temporally modulated backgrounds: A possible physiological explanation based upon a recent computational model. *Vis. Neurosci.*, 15: 957–967.
- Hood, D. C., Graham, N., von Wiegand, T. E. & Chase, V. M. (1997). Probed-sinewave paradigm: A test of models of light-adaptation dynamics. *Vis. Res.*, 37: 1177–1191.
- Landy, M. & Graham, N. (2003). Visual perception of texture. In L. M. Chalupa & J. S. Werner (Eds.) *The Visual Neurosciences*. pp. 1106–1118. MIT Press: Cambridge, MA.
- Lennie, P. (1998). Single units and visual cortical organization. *Percept.*, 27: 889–935.
- Snippe, H. P., Poot, L. & van Hateren, J. H. (2000). A temporal model for early vision that explains detection thresholds for light pulses on flickering backgrounds. *Vis. Neurosci.*, 17: 449–462.
- Snippe, H. P., Poot, L. & van Hateren, J. H. (2004). Asymmetric dynamics of adaptation after onset and offset of flicker. *J. Vis.*, 4: 1–12.
- Wilson, H. R. (1997). A neural model of foveal light adaptation and afterimage formation. *Vis. Neurosci.*, 14: 403–423.
- Wilson, H. R. & Humanski, R. (1993). Spatial frequency adaptation and contrast gain control. *Vis. Res.*, 33: 1133–1149.
- Wolfson, S. S. & Graham, N. (2000). Exploring the dynamics of light adaptation: The effects varying the flickering backgrounds duration in the probed-sinewave paradigm. *Vis. Res.*, 40: 2277–2289.
- Wolfson, S. S. & Graham, N. (2001a). Comparing increment and decrement probes in the probed-sinewave paradigm. *Vis. Res.*, 41: 1119–1131.
- Wolfson, S. S. & Graham, N. (2001b). Processing in the probed-sinewave paradigm is likely retinal. *Vis. Neurosci.*, 18: 1003–1010.
- Wolfson, S. S. & Graham, N. (2005a). Element-arrangement textures in multiple objective tasks. *Spatial Vis.*, 18: 209–226.
- Wolfson, S. S. & Graham, N. (2005b). Dynamics of contrast-gain controls in pattern vision. *Vis. Sci. Soc.*, abstract #762
- Wu, S., Burns, S. A., Elsner, A. E., Eskew Jr., R. T. & He, J. (1997). Rapid sensitivity changes on flickering backgrounds: Tests of models of light adaptation. *J. Opt. Soc. Am. A*, 14: 2367–2378.