

TUNING AND BINDING IN COMPOUND ADAPTATION

David McGovern, Sarah Hancock and Jonathan W. Peirce

Nottingham Visual Neuroscience, School of Psychology, University of Nottingham

Abstract

The presence of neural mechanisms selective to plaid patterns has been shown using compound adaptation. We used this method to investigate the constraints the visual system employs when combining V1 outputs. Participants compared the contrast of a probe in one visual hemi-field to a fixed reference probe in the other, following adaptation to a plaid or its components respectively. Experiment 1 showed that compound adaptation was tightly tuned to the spatial frequency of the adaptor relative to the test. Experiment 2 demonstrated that adaptation was strongest when the components shared the same spatial frequency and decreased markedly when they differed. Experiment 3 varied the temporal onset of the components comprising the plaid adaptor and found a non-linear decrease in adaptation; small temporal separations of the components caused large decreases in adaptation. These findings show that small component differences may impact on the binding of these components into a single plaid texture.

A common function of many neurons throughout mammalian cortex is to combine the outputs of other neurons and provide a transformed representation of that combined signal. In the visual system this is usually achieved through a feed-forward network, where the outputs from a previous level of analysis are combined leading to more detailed representations of a stimulus in progressively higher visual areas (Felleman & Van Essen, 1991). For instance, the outputs of the retinal ganglion cells are combined to form the centre-surround antagonistic receptive fields found in the lateral geniculate nucleus (LGN). Similarly, the projections from the retinal ganglion cells are passed through the LGN to form the elongated receptive fields of V1.

Rather little is known about the combinatorial processes beyond V1 and what type of receptive field may result. It seems likely, however, that the visual system employs a similar strategy to that used at earlier levels of analysis and pools the local orientation signals from V1. Potentially, responses from V1 with different preferred orientations could be summed to form contours of various curvatures. Certainly it has been shown that V2 neurons encode combinations of different orientations within their receptive field (Anzai, Peng & Van Essen, 2007) and that neurons in V4 appear to be sensitive to the presence of curved contours on their receptive fields (e.g Gallant et. al. 1993). Additionally psychophysical adaptation aftereffects have shown that the visual system demonstrates selectivity to combinations of curved contours that cannot be explained by local tilt-aftereffects (Hancock & Peirce, 2008). For overlapping V1 receptive fields a similar procedure could be utilised to create a mechanism selective for a plaid pattern. Although physiological studies have yet to report the existence of

neurons that respond selectively to plaid patterns, psychophysical evidence suggests their presence (Peirce & Taylor 2006; McGovern & Peirce, in press).

There is a problem, however, with the idea that at some stage beyond V1, local orientation signals are combined to form compound detectors, which is that there are too many pairwise combinations of these cells to be represented. If, for example, the set of V1 neurons were governed by N parameters (e.g spatial frequency, orientation bandwidths) and contained M cells, then the set of pairwise combinations has $2 \times N$ parameters and M^2 cells. This problem is accentuated by the fact that V1 is the largest of the visual areas. Thus, it appears that the visual system must, in some way, constrain which outputs are combined beyond V1.

Here we address three possible ways in which the visual system might ameliorate this “parameter explosion”. The first possible solution is for a compound detector to be invariant to some spatial feature of the stimulus, which is encoded at a previous processing stage. For instance a plaid-detector might respond to any plaid pattern irrespective of its spatial frequency, such that spatial frequency, as a parameter of the plaid, is removed (Experiment 1). Alternatively, a compound detector might selectively choose which V1 signals to combine, that is a plaid mechanism might only respond to plaids comprised of components with similar features (Experiment 2). Finally, a plaid mechanism may be sensitive to small temporal changes, such that they only respond when the plaid components are presented in synchrony (Experiment 3).

Methods

All experiments used a ‘compound adaptation’ paradigm, designed to measure adaptation to a compound stimulus beyond that predicted by adaptation to its constituent parts. Two patches on opposite sides of the visual field are adapted simultaneously – one to a plaid stimulus consisting of two gratings presented together (compound field), and one to the same two grating stimuli presented in isolation, alternating every second (component field). A test stimulus is then presented in both adapted locations; one at the same retinal location that it had itself been adapted (the test probe) with one in the opposite location (the reference probe). The reference probe held a fixed Michelson contrast of 42%. The participant is required to judge which side had the higher apparent contrast with the point of subjective equality derived through a 1-up, 1-down staircase procedure. Each probe type (plaid, component A and component B) was tested in a separate staircase. As the same component gratings are presented in both adapting locations any aftereffect due to adaptation to the components alone should be equal on both sides. Therefore, any residual difference in the adaptation effect between the two sides must be due to adaptation to the compound as a whole.

Plaids were constructed from the linear combination of two luminance-modulated sinusoidal gratings at oblique angles, $\pm 45^\circ$ from vertical. Each grating contributed equal contrast to the plaid (50% Michelson). The plaid stimulus alternated with a blank field every second to equate the exposure time for each of the two gratings in both hemi-fields. All stimuli were presented in a Gaussian envelope with a

standard deviation of 0.5° visual angle (such that the stimulus had a diameter of 8° at the point where it fell below 1% contrast). The spatial phase of the stimulus was randomly jittered (every 200ms) across time to prevent retinal afterimages.

Results-Experiment 1

In order to avoid the parameter explosion involved in combining V1 outputs, the visual system may act to discard some stimulus features processed in previous visual areas. Specifically, the plaid-selective mechanisms in question could respond to any plaid irrespective of its spatial frequency. This would greatly reduce the neural machinery required to process compound stimuli. To test this hypothesis, Experiment 1 obtained spatial frequency tuning functions for plaid patterns by adapting participants to stimuli of a fixed spatial frequency who were subsequently tested with a probe that could take one of five spatial frequency values depending on the trial.

Figure 1 shows mean data for four participants collected with adapting stimuli of 1.26 c/deg. Peak adaptation effects of almost 3db are observed in trials where adaptor and probe share a common spatial frequency, with adaptation dropping to 1db for probes at either extreme of the function. These results indicate that the mechanisms responsible for processing plaid patterns are strongly tuned for spatial frequency and are clearly at the odds with the notion of spatial frequency invariant conjunction detectors. Thus, conjunction detectors must employ another strategy in pooling local orientation signals from V1.

Experiment 2

Experiment 1 demonstrated that compound adaptation is tightly tuned for the spatial frequency of the test probe, implying that the visual system does not completely discard the spatial frequency content of plaid patterns to circumvent the combinatorial explosion involved in mid-level feature integration. A second way to ameliorate the parameter explosion is for conjunction detectors to respond only to combinations of stimuli that share similar features, which may act as a cue to the components belonging to the same object. To test this hypothesis, Experiment 2 examined the effect of varying the spatial frequencies of the components within the compound relative to each other. Here participants were adapted to plaids which, depending on the trial, had components of the same (low, medium and high conditions) or different (1 octave and 3 octave conditions) spatial frequencies. In this experiment the test probe was always the same as the adaptor.

Figure 1B shows mean data for four participants. Greatest adaptation effects were found when the plaid was comprised of components with a common spatial frequency. Although a moderate decrease is observed in the high spatial frequency condition for the mean data, this decrease is within the margin of error. When the plaid components are separated by just one octave a large decrease in plaid adaptation is observed, with a further sharp decrease with a three octave separation. Thus, not all pairwise combinations of gratings lead to large compound adaptation effects. When

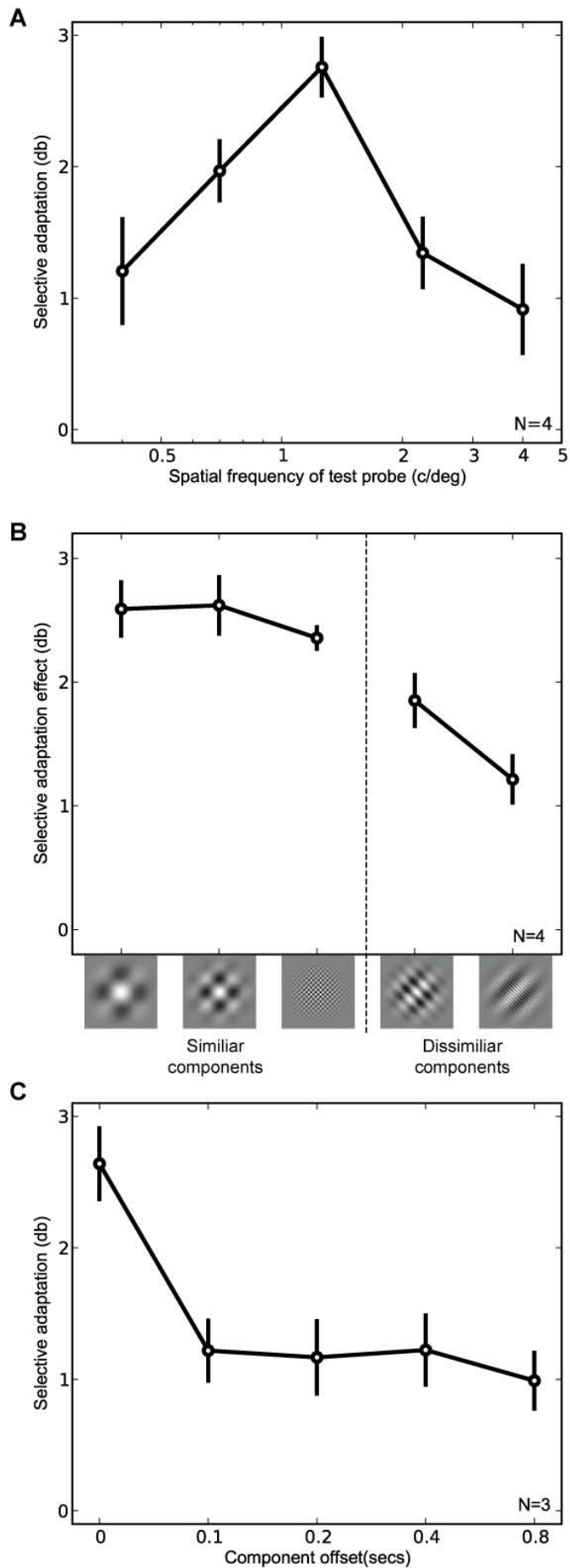


Figure 1: (A) Plaid-selective adaptation was strongly tuned for spatial frequency. (B) Combinations of dissimilar components led to weaker adaptive effects. (C) Separating the components in time led to a non-linear decrease in adaptation. See text for further details.

the component gratings that comprise the plaid are substantially different a reduced adaptive effect is observed.

Experiment 3

Experiments 1 & 2 examined two potential strategies that the visual system may employ to limit the amount of pairwise V1 outputs represented in later processing stages. Additionally, the visual system might only respond to a plaid pattern if its components are presented in synchrony. To test this hypothesis, Experiment 3 investigated how varying the onset of the components comprising the plaid pattern impacted on plaid-selective adaptation. In this experiment the temporal phase of the components was offset taking one of six values between 0 seconds (full coherent plaid) and 1 second (components completely out of synchrony). For values in between these extremes, the components over-lapped for a portion of the stimulus cycle determined by the offset value.

Figure 1C shows mean data for three participants. The results show a non-linear decrease in plaid-selective adaptation with increasing temporal separation of the components. When the components temporally coincide, a 2.5 decibel adaptation effect is observed. Small temporal separations (100 ms) lead to large decreases in adaptation with further temporal separations leading to much smaller decreases. This non-linear decrease suggests that the mechanisms mediating plaid-selective adaptation are sensitive to the temporal properties of the components comprising the plaid pattern, which could help in dissociating sensory information from different objects.

Discussion

It is well known that orientated receptive fields of V1 neurons are generated by summing appropriately selected retinal input, relayed through the LGN. Much less is known about how the subsequent levels combine the outputs from V1. Stimulus selective aftereffects reveal that the outputs of overlapping receptive fields can be combined to represent a plaid pattern. However, due to the large amount of different pairwise combinations of V1 signals, it is highly improbable that all conjunctions are represented. Therefore, in order for the visual system to form an economical representation of the world, it must, in some way, constrain which outputs are represented as compounds beyond V1.

To achieve this, a putative compound detector may be invariant to some aspect of the stimulus, say spatial frequency, which has been processed at an earlier level of analysis. By removing spatial frequency as a parameter of the plaid, the visual system could transmit this information more efficiently, helping to reduce redundancy (Barlow, 1961). This, however, was shown not to be the case. The results from Experiment 1 clearly show that plaid-selective adaptation is tightly tuned for spatial frequency, with the largest aftereffect observed when the plaid probe shared a common spatial frequency with the adaptor. Thus, spatial frequency information appears to be retained by the compound detector, which suggests that these

mechanisms must employ an alternative strategy in pooling local orientation signals from V1.

An alternative solution to this problem is for a compound detector to respond to a plaid only when the features of the components are similar. This approach would greatly reduce the amount of neural machinery required to process plaid patterns, as mid-level neurons would need only to respond to a subset of the potential combinations from V1 to represent a plaid. The results from Experiment 2 support this interpretation, showing that maximal adaptation is observed when the components comprising the plaid have the same spatial frequency. This adaptation effect decreases significantly when the components are separated by just one octave and is further diminished with a three octave separation. These results suggest that only outputs representing stimuli with similar low-level features are initially combined after V1.

This selective combination of outputs could play an important role in the segmentation of a scene. Given that the retina simply projects a two-dimensional signal representing points of light that land on it, it follows that the visual system must be equipped with cortical mechanisms capable of generating different segmented representations of an image. It is also likely that the visual system exploits temporal cues in the environment to aid this process. The results from Experiment 3 demonstrate that the underlying mechanisms mediating plaid-selective adaptation are very sensitive to the temporal properties of the plaid stimulus. That is, plaid-selective adaptation decreased significantly when the temporal phases of the components comprising the pattern were offset by just 100ms. Taken together, the spatial and temporal non-linearities inherent in combining V1 outputs to form a coherent plaid texture suggest an important role for the underlying mechanisms in segmentation processes.

References

- Anzai, A., Peng, X., & Van Essen, D. C. (2007). Neurons in monkey visual area V2 encode combinations of orientations. *Nat Neurosci*, *10*(10), 1313-1321.
- Barlow, H.B. (1961). The coding of sensory messages. In W.H. Thorpe and O.L. Zangwill (Eds), *Current Problems in Animal Behavior*, Cambridge: Cambridge University Press. pp. 331–360.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cereb. Cortex*, *1*(1), 1-47.
- Gallant, J. L., Braun, J., & Van Essen, D. C. (1993). Selectivity for Polar, Hyperbolic, and Cartesian Gratings in Macaque Visual-Cortex. *Science*, *259*(5091), 100-103.
- Hancock, S., & Peirce, J. W. (2008). Selective mechanisms for simple contours revealed by compound adaptation. *Journal of Vision*, *8*(7), 1-10.
- McGovern, D.P., & Peirce, J.W. (in press). The spatial characteristics of plaid-form-selective mechanisms. *Vis Res*.
- Peirce, J. W., & Taylor, L. J. (2006). Selective mechanisms for complex visual patterns revealed by adaptation. *Neuroscience*, *141*, 15-18.