Spatial and Temporal Interactions between Shape Representations in Human Vision

SPATIAL AND TEMPORAL INTERACTIONS BETWEEN SHAPE REPRESENTATIONS IN HUMAN VISION

BY

MICHAEL SLUGOCKI, B.Sc.

A THESIS SUBMITTED TO THE

DEPARTMENT OF PSYCHOLOGY, NEUROSCIENCE & BEHAVIOUR

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

McMaster University

DOCTOR OF PHILOSOPHY (2019)   McMaster University

(Psychology, Neuroscience & Behaviour)   Hamilton, Ontario, Canada

TITLE: Spatial and Temporal Interactions between Shape Rep-
                    resentations in Human Vision

AUTHOR: Michael Slugocki

                    B.Sc., (Psychology, Neuroscience & Behaviour)

                    McMaster University, Hamilton, Canada

SUPERVISORS: Dr. Patrick J. Bennett

                    Dr. Allison B. Sekuler

NUMBER OF PAGES: xxvi, 227

Lay Abstract

The extraordinary ability of humans to identify objects is due in part to the ability of the visual system to accurately represent the overall shape of an object's outline. Human perception of outline shape has been well studied using single-object displays, but natural scenes typically contain multiple objects and therefore multiple shapes. In general, the perception of one part of a visual scene depends on other parts of the scene, and therefore it is likely that the perception of an object's shape also depends on the other objects in a scene. Therefore, the goal of this thesis was to examine how the perception of shape is affected when multiple objects are viewed. We found that human sensitivity for changes in shape is impaired when multiple objects appear close to one another in space and/or time. Furthermore, the way that shape perception is affected by multiple objects offers important clues about the features that the visual system uses to represent shape.

Abstract

The human visual system has the remarkable capacity to transform spatio-temporal patterns of light into structured units of perception. Much research has focused on how the visual system integrates information around the perimeter of closed contours to form the perception of shape. This dissertation extends previous work by investigating how the perception of curvature along closed-contour shapes is affected by the presence of additional shapes that appear close to the target shape in space and/or time.

Chapter 2 examined the ability of shape mechanisms at representing low frequency curvature in the presence of a higher frequency component along contours in multi-shape displays. We found that additions of high amplitude, high frequency curvature along a contour path can modulate the strength of interaction observed between shapes, and thus attenuates the contribution of low frequency components in interactions between neighbouring contours. Chapter 3 examined what curvature features are of importance in modulating phase dependent interactions between shapes. Results revealed that phase-dependent masking does not depend on curvature frequency, but is related to sensitivity for phase shifts in isolated contours, and is affected by both positive and negative curvature extrema. Computational simulations aimed at modelling the population responses evoked in intermediate shape processing areas

(i.e., V4) suggest sensitivity to shifts in phase of shapes is not well captured by such a population code, and therefore alternative explanations are required. Chapter 4 examined how sensitivity to curvature deformations along the contour of a closed shape changes as a function of polar angle, angular frequency, and spatial uncertainty. Results show that human observers are, at first approximation, uniformly sensitivity to curvature deformations across all polar angles tested, and this result holds despite changes in angular frequency and spatial uncertainty. Chapter 5 examined whether the strength of spatial masking between shapes is affected by the presentation of a temporal mask. Our results demonstrate that a temporal mask affected spatial masking only when it preceded the target-mask stimulus by 130-180 ms. Furthermore, the effects of a temporal mask on spatial masking are approximately additive, suggesting that separate components contribute to spatial and temporal interactions between shapes.

Preface

I dedicate this work to my brother, and best friend, Christopher.

To my sweetest Catherine.

And to my loving parents, Henryka and Zbigniew Slugocki.

This dissertation comprises 6 chapters and is written in the "sandwich" thesis format. Chapter 1 provides a general overview of the current literature on shape perception and the aim of this dissertation. Chapters 2-5 describe the research in a journal article format. Chapter 6 discusses the implications, limitations, and future avenues of study for this research.

Research in Chapter 2 was conducted in 2014 and will be submitted to Vision Research. Research in Chapter 3 was conducted in 2015 and was published in Vision Research in 2019. Research in Chapter 4 was conducted in 2017 and will be submitted to Vision Research. Research in Chapter 5 was conducted in 2016 and has been accepted for publication in Journal of Vision for 2019.

The research described in this dissertation was conducted in collaboration with my supervisors, Dr. Patrick J. Bennett and Dr. Allison B. Sekuler. All major components involved in conducting this research were carried out by the principle author,

Michael Slugocki. These responsibilities include, but are not limited to: development of research design and methods, programming and validation of experimental setup, data analysis, and writing first version of manuscripts. Subsequent versions of manuscripts were revised in collaboration with the co-authors of each paper. Donna Waxman, the lab coordinator, was responsible for recruiting and testing observers who were non-authors, and whose data are featured in each study.

Acknowledgements

I am tremendously grateful to my supervisors, Dr. Patrick J. Bennett and Dr. Allison B. Sekuler for providing me with the opportunity to conduct research at the highest scholarly level. My deepest appreciation to them both for imparting their passion for quality science, wisdom in studying human vision, and continual encouragement to push the boundaries of human knowledge. Also, thank you to Dr. David I. Shore for being a member on my thesis committee, and his insightful feedback during our annual meetings.

Thank you to my friend, and lab research assistant, Donna Waxman for her tireless charisma and amiable disposition that always serves to bring out the best in others. Also, to my colleagues in the Vision Lab, and to the Department of Psychology, Neuroscience & Behaviour for creating an interesting place of work and study.

Many thanks to all of my friends and family for their unconditional love and encouragement. For my partner, Catherine, whose radiance continues to fuel my heart throughout the toughest of times. A very special thank you to my brother, Christopher, whom I always look up to as a mentor in life. It is his enthusiasm and devotion to science and philosophy that inspired me to pursue academia. And, to my loving parents to whom I am forever grateful. It is their continual sacrifice, effort, love, and support throughout my life that has made opportunities like this possible.

Contents

List of Tables

List of Figures

xxv

Chapter 1

Introduction

1.0.1 Local filters to global percepts along the visual hierarchy

The ability of the visual system to accurately encode the form of objects is critical in guiding purposeful interactions with those objects and the environment. In order for objects to be represented by biologically plausible neural codes, the visual system must overcome the computational challenge of parsing a visual scene into discrete elements, and grouping those elements to form a coherent set of features that themselves are combined to form larger structured units. Gestalt psychologists of the early 20th century first described a set of principles which the visual system relies on to segment and group elements in visual space (Koffka, 1922; Köhler, 1938; Wagemans *et al.*, 2012a,b; Wertheimer, 1912, 1923). Grouping principles, such as those of proximity and good continuation, serve to organize the visual array into units that correspond closely to surfaces and objects in the visual scene, and that form the foundation of perception (Koffka, 1922; Köhler, 1938; Wagemans *et al.*, 2012a,b; Wertheimer, 1912,

1923).

These early descriptive approaches at modelling human vision led to tremendous advances in our understanding of perception, and revealed an important characteristic of vision: changes to localized areas in a stimulus can profoundly change how the stimulus as a whole is perceived (Wagemans *et al.*, 2012a,b). However, neurons within early stages of visual processing (V1) have receptive fields that are highly localized, and selectively tuned to respond to only simple features, such as orientation and spatial frequency (Bredfeldt and Ringach, 2002; Foster *et al.*, 1985; Gilbert, 1977; Hubel and Wiesel, 1959, 1968; Hubel *et al.*, 1978; Olshausen and Field, 2005; Ringach *et al.*, 2002; Rust *et al.*, 2005). As a consequence of the limited visual input received by each neuron in V1, defined by a neuron's receptive field, individual V1 neurons are not able to determine if an element within its receptive field is presented in isolation or is part of an extended curve (Loffler, 2015). Visual information about an object typically is distributed across extended regions of a scene, and therefore there is a strong functional need for these early localized filters, and those in subsequent visual areas, to pool information across space in order to accurately represent increasingly complex features of an object (Wilson and Wilkinson, 2015).

The first step in signal integration likely originates at the level of V1 via long-range lateral interactions between cells. Evidence from anatomical and neurophysiological studies of different mammalian species demonstrate that long-range horizontal projections from orientation columns in V1 typically converge onto columns that share a similar orientation preference (Bosking *et al.*, 1997; Das and Gilbert, 1995; Gilbert, 1992; Gilbert and Wiesel, 1979, 1983, 1989; Rockland and Lund, 1983; Stettler *et al.*, 2002; Ts'o and Gilbert, 1988; Ts'o *et al.*, 1986). These findings suggest that the

arrangement of V1 neurons may facilitate the grouping of similarly oriented visual elements across different spatial scales (Field *et al.*, 1993). Examination of edge properties within images of natural scenes reveal the importance of adopting such a connectionist architecture, as local edge elements that are separated in distance, but belong to the same contour, tend to be aligned along a linear or co-linear path (Elder and Goldberg, 2002; Geisler *et al.*, 2001). Thus, at early stages of visual processing, the visual system is able to filter images to extract and represent more structured image features, such as extended curvature segments, that can be combined to generate more complex representations of form. Orientation linking and combining elementary features to represent the outline of objects at low levels in the visual hierarchy form the first stages of shape perception (Hess *et al.*, 2014).

Beyond V1, there continues to be an increase in the computational complexity of how information is represented within the visual cortex (Loffler, 2015), as well as an increase in the receptive field size of cells at higher cortical areas (Dumoulin and Wandell, 2008; Smith *et al.*, 2001). At V2, cells display selective tuning for image features, such as angles and arcs, that require the integration of information across orientation selective cells in V1 (Hegdé and Essen, 2000; Ito and Komatsu, 2004; Kobatake and Tanaka, 1994; Sincich and Horton, 2005). Response selectivity in V4 continues to change as cells respond only modestly to stimuli that lie along a Cartesian coordinate system, and instead prefer a polar-based coding scheme (Dumoulin and Hess, 2007; Gallant *et al.*, 1992, 1996; Pasupathy and Connor, 1999, 2001, 2002; Wilkinson *et al.*, 2000). Neurophysiological studies in macaques suggest the polar-based coding scheme allows V4 neurons to generate a population code that represents curvature along the outline of a closed contour relative to the object center (Gustavsen and

Gallant, 2003; Pasupathy and Connor, 2001, 2002). Human lateral occipital complex (LOC) demonstrates selective activation to object identity, and is invariant to local properties of shape, such as the contours that define the outline of the object being viewed (Grill-Spector *et al.*, 1999; Kourtzi and Kanwisher, 2000, 2001). Other extrastriate areas presumably use these representations to construct specialized representations of specific object categories, such as faces or body parts (Downing *et al.*, 2001; Grill-Spector *et al.*, 2004). Ultimately, convergence and summation of visual information along stages of visual processing produce structured representations that are more useful for discriminating and identifying objects than are the outputs from earlier cortical areas.

Along all stages of visual processing, neuronal responses typically depend on the stimulus falling within a neuron's receptive field as well as the stimulus falling outside the classical receptive field. This type of contextual modulation of a neuronal response depends on interactions among neurons that are situated close together, as well as interactions arising from long-range connections within a cortical area (e.g., long-range horizontal connections in V1) and between visual areas via feedforward and recurrent connections (Lamme and Roelfsema, 2000). Therefore, there is no single locus that integrates all visual inputs into a coherent percept, but rather it is better described by a network of interactions among neurons distributed in different visual cortical areas, perform unique image analyses at a variety of spatial scales (Van Essen and Gallant, 1994; Wilson and Wilkinson, 2015). The consequence is a biological system that systematically transforms a set of visual inputs, through linear and non-linear transformations (Brincat and Connor, 2004), into a set of structured elements such as those first described by early Gestalt Psychologists as the fundamental principles

of perception.

1.0.2   Radial frequency contours and shape perception

The detection of curvature maxima plays a fundamental role in the perception of shape (Attneave, 1954). Information regarding the identity and overall form of an object appears to be concentrated along points of maximum curvature, such that object recognition remains intact even when these points are connected via straight edge segments (Attneave, 1954). Studies examining the ability of the human visual system to perceive curvature maxima, or more broadly curvature, have used made use of a variety of stimuli that include angles (Chen and Levi, 1996; Heeley and Buchanan-Smith, 1996; Kennedy *et al.*, 2006, 2008; Regan *et al.*, 1996), sinusoidal contours (Gheorghiu and Kingdom, 2006, 2007; Prins *et al.*, 2007; Tyler, 1973), and curved arcs (Wilson and Richards, 1989, 1992). Although humans are sensitive to changes in curvature along the outline of an open contour, studies have demonstrated a perceptual advantage for detecting closed contour stimuli embedded in noise fields (Kovács and Julesz, 1993). Furthermore, humans are are extremely sensitive to changes in curvature along the contour of a circle and similar smooth, closed-contour stimuli (Baldwin *et al.*, 2016; Green *et al.*, 2017, 2018a; Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann and Kingdom, 2017; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998). These findings have led to a tremendous effort in recent years to better understand how curvature of smooth, closed contours is represented by the visual system from both a physiological (Gallant *et al.*, 1996; Pasupathy and Connor, 1999, 2001, 2002) and behavioural (Baldwin *et al.*, 2016; Bell and Badcock, 2009; Bell *et al.*, 2008, 2009; Green *et al.*, 2017, 2018a,b,c; Habak *et al.*, 2004, 2006; Hess *et al.*, 1999; Jeffrey

et al., 2002; Kurki *et al.*, 2014; Loffler *et al.*, 2003; Schmidtmann and Kingdom, 2017; Schmidtmann *et al.*, 2012; Wilder *et al.*, 2018; Wilkinson *et al.*, 1998) viewpoint.

In an effort to better understand how the human visual system represents smooth deformations along closed contour shapes, Wilkinson *et al.* (1998) developed a stimulus class known as Radial frequency (RF) contours (Figure 1.1). These stimuli are useful for studying shapes, as they serve as a way of parameterically controlling for the generation of shapes through the sinusoidal modulation of the radius of a circle according to:

$$r(\theta) = \bar{r}(1 + A\sin(\omega\theta + \phi)) \tag{1.1}$$

where $\bar{r}$ is the mean radius of the contour, A defines the amplitude of modulation as a proportion of the radius of the circle, ω the radial frequency number, and ϕ the angular phase. By varying amplitude, radial frequency, and phase, whole families of smooth contour shapes can be generated where the precise quantitative relation between shapes are well defined. Manipulation of these parameters, in addition to combining different RF contours, can generate a variety of geometric shapes, some of which can mirror natural complex forms such as the outline of a human head (Anderson and Wilson, 2005; Wilson and Wilkinson, 2002; Wilson *et al.*, 2000, 2002, see Figure 1.2), or the torso of animals (Wilson and Wilkinson, 2002). While these patterns are useful for their mathematical properties and biological relevance, it is important to acknowledge that these patterns represent only a small subset of all possible planar shapes (Schmidtmann and Fruend, 2019).

Wilkinson *et al.* (1998) originally used RF contours to examine human sensitivity to deviations from circularity using classical psychophysical methods. Wilkinson *et al.*
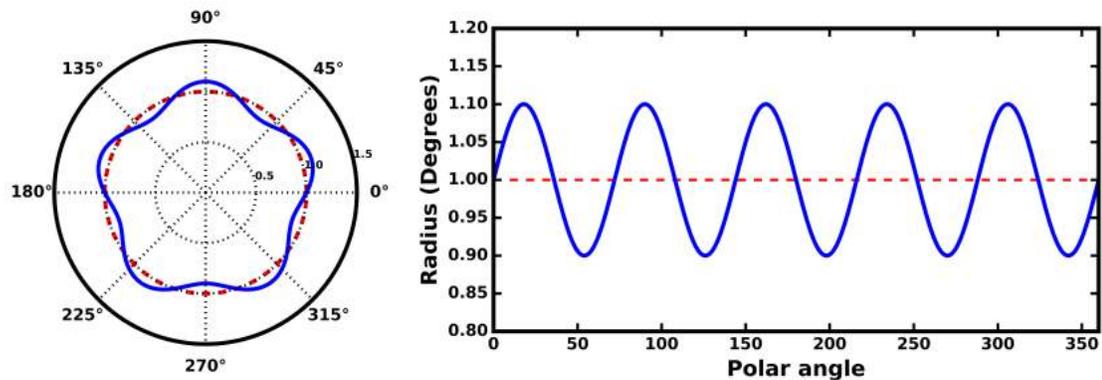
Figure 1.1: Example of a Radial Frequency (RF) contour that contains 5 cycles of modulation per 2π (blue) compared to a perfect circle (red). The left figure displays plots each stimulus along polar coordinates demonstrating how the application of radial modulations along the circumference of a circle can produce a smooth geometric closed contour that resembles common shapes. The right figure plots the magnitude of radial modulation applied along each polar angle, here using a base radius of 1 visual degree. The derivation of RF contours by Wilkinson *et al.* (1998) has allowed researchers to better understand how the human visual system is able to perceive curvature along the outline of a closed contour.
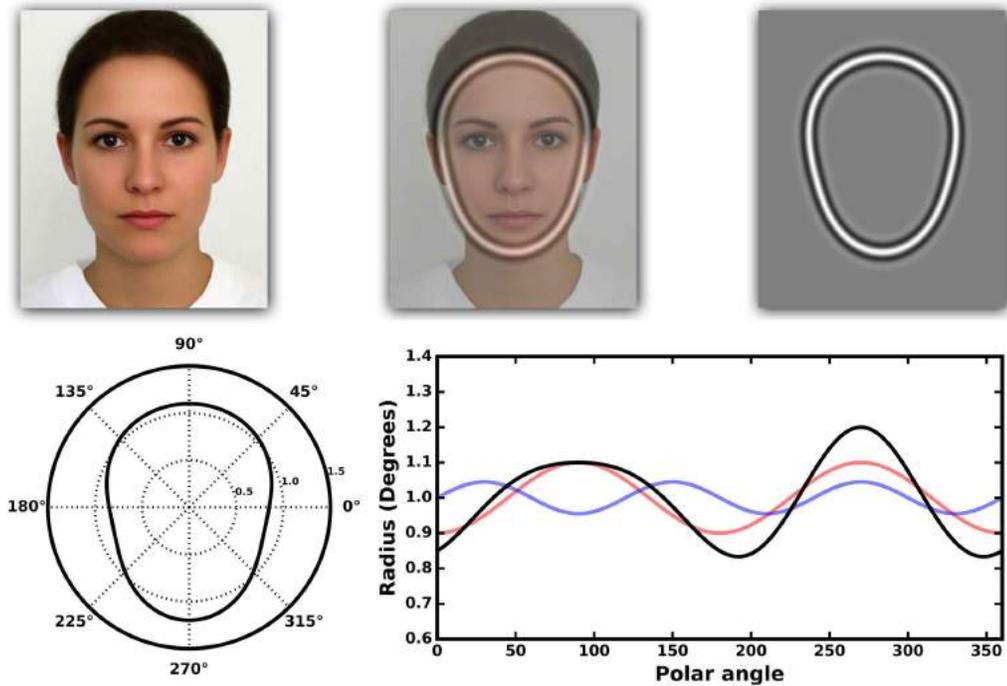
Figure 1.2: *Upper panel:* Example of how the general outline of a face can be represented by two radial frequency components (RF2 + RF3), with the addition of more components allowing for an increase in the accuracy of representation of the outline. *Lower panel:* The left figure displays a polar plot of the compound RF contour featured in the upper panel. The right figure plots the magnitude of radial modulation of each RF component, RF2 (red) and RF3 (blue), and the compound RF contour that results to form the outline of a head. By manipulating parameters that comprise each RF component, and adding different RF components together, a variety of common geometric forms can be generated allowing researchers precise control over the parametric properties of different natural shapes (Wilson and Wilkinson, 2002; Wilson *et al.*, 2002).

found that human observers exhibit hypersensitivity to curvature in such conditions: in other words, under optimal conditions the amount of curvature that was needed to discriminate between a perfect circle and a deformed circle was so small that the maximum difference between the retinal images of the two contours was smaller than the distance between adjacent foveal cones (Westheimer, 1975). Thresholds also differed as a function of the number of modulations along the circumference of the base circle (i.e., the radial frequency (RF) of the contour), where best performance is observed for RFs ≥ 3 (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998). Such remarkable sensitivity at detecting curvature along RF contours usually has been attributed to the global integration of curvature information for RF contours between approximately 3-8 cycles per circumference (Green *et al.*, 2017, 2018a; Jeffrey *et al.*, 2002; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998).

The degree to which local signals are integrated by a global operator is assessed using experimental paradigms that restrict the number of visible deformations to only a fraction of the entire contour (Baldwin *et al.*, 2016; Green *et al.*, 2017, 2018a; Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). While methodological details differ across studies, these experiments typically measure how sensitivity to contour curvature depends on how the signal – i.e., curvature – is distributed along the length of a contour. The rate at which performance changes as a function of signal strength can be used to distinguish among various processing strategies, such as probability summation or true summation of local measures of curvature. For low radial frequency contours – i.e., those that contain less than 8 - 10 cycles of modulation – detection thresholds improve with increasing number of RF cycles at a

rate that that is faster than the rate predicted by probability summation alone (Green *et al.*, 2017, 2018a; Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; but see Baldwin *et al.*, 2016 for an alternative explanation). This finding has been taken as evidence that low radial frequency curvature is detected by mechanisms that efficiently combine local curvature information across the entirety of a shape (Green *et al.*, 2017, 2018a; Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). In contrast, detection thresholds for high radial frequencies – i.e., those above 8 - 10 cycles of modulation – fall with increasing number of curvature cycles at a rate that is consistent with probability summation (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012), which is consistent with the ideal that local computations limit the detectability of high frequency RF contours.

The fact that curvature summation differs between low and high radial frequencies has led some researchers to argue for the existence of RF-selective shape channels in human vision (Hess *et al.*, 1999; Jeffrey *et al.*, 2002; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998). Studies using subthreshold summation (Bell *et al.*, 2009), adaptation (Bell and Badcock, 2009; Bell *et al.*, 2008), and masking (Habak *et al.*, 2004, 2006) have provided evidence in support of this proposition; however, the most compelling evidence comes from studies using compound RF contours (Bell and Badcock, 2009; Bell *et al.*, 2007), which is a single contour that contains curvature at several radial frequencies. Several studies have compared curvature detection thresholds measured with compound RF contours to thresholds for the individual RF components and found that detection threshold for the compound contour is significantly higher than thresholds for the individual RF components provided that the components have similar frequencies (Bell and Badcock, 2009; Bell *et al.*, 2007).

However, if the frequency of components combined to form the compound are disparate, little to no interactions between components are observed (Bell and Badcock, 2009; Bell *et al.*, 2007). The results from these studies support the existence of multiple, independent shape channels that are tuned to a narrow frequency of curvature occurring along the outline of a shape (Bell and Badcock, 2009; Bell *et al.*, 2007).

Findings from physiological and behavioural investigations of shape perception, many of which utilized RF contours as stimuli, has led to the development of computational models of how the visual system represents shape information (Kempgens *et al.*, 2013; Poirier and Wilson, 2006; Schmidtmann and Kingdom, 2017). Given that V4 employs a population code to represent curvature relative to the center of an object (Carlson *et al.*, 2011; Pasupathy and Connor, 2001, 2002), similar population encoding strategies have been used to model how the human visual system encodes stimuli such as RF contours (Kempgens *et al.*, 2013; Poirier and Wilson, 2006). A common characteristic shared by these models includes using a polar-based coding scheme to represent curvature maxima relative to an object's center (Kempgens *et al.*, 2013; Poirier and Wilson, 2006), which allows for subsets of units within the population to encode both the location and magnitude of deformities along the contour of a shape. The pattern of activation across the population can then be used to differentiate between similar shapes, or to reconstruct a representation of the original shape (Kempgens *et al.*, 2013; Poirier and Wilson, 2006). Such computational models offer a link between physiological investigations of shape-encoding neurons in intermediate visual areas to behavioural measures of performance.

To summarize, RF contours have been extensively used in the study of human

shape perception (Loffler, 2008, 2015). Observers are extremely sensitive to curvature, (Habak *et al.*, 2004; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998), which follows as a consequence of global processes integrating local curvature information across the entirety of a shape (Bell and Badcock, 2009; Bell *et al.*, 2008, 2009; Green *et al.*, 2017, 2018a; Jeffrey *et al.*, 2002; Kurki *et al.*, 2014; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998). To effectively and efficiently encode shapes, computational models suggest that the visual system represents shapes in a multidimensional shape space, whereby shape features, such as curvature maxima, are represented along polar coordinates relative to an object's center (Carlson *et al.*, 2011; Dickinson *et al.*, 2013, 2015, 2018; Kempgens *et al.*, 2013; Pasupathy and Connor, 2001, 2002; Poirier and Wilson, 2006).

1.0.3   Visual masking as a tool to study vision

Masking is a commonly used psychophysical tool that has a long history of use in vision science to better understand visual processes. In general, masking refers to a reduction in the visibility of a target stimulus that can be attributed to the presentation of a second stimulus, referred to as the mask, that occurs within a proximate spatio-temporal window relative to the onset of the target object (Breitmeyer and Ogmen, 2000; Enns and Di Lollo, 2000; Lev and Polat, 2015). Masking that results from spatial interactions between a target and mask often is categorized into two distinct subtypes: pattern masking and lateral masking (Enns and Di Lollo, 2000; Lev and Polat, 2015). Pattern masking refers masking that occurs when the target and mask share the same retinal location, while lateral masking is produced by a mask that is positioned near a target, but there is no overlap between the spatial location

of both objects (Lev and Polat, 2015). Regardless of the spatial masking paradigm employed, both subtypes serve as useful tools to probe steady state properties of spatial-processing mechanisms (Breitmeyer and Ogmen, 2000).

Spatial domain

Within the spatial domain, previous studies have made extensive use of pattern masks to investigate the existence and bandwidth of visual spatial channels (Graham and Nachmias, 1971). For example, Graham and Nachmias (1971) measured detection thresholds for stimuli consisting of two superimposed sinuisoidal gratings and found that the variation in detection thresholds – i.e., the amount of masking – depended on the relative spatial frequencies of the two gratings. Specifically, masking was high when the two spatial frequencies were similar, and near zero when the frequencies differed by about a factor of three. This pattern of results is consistent with the idea that sinusoidal gratings are detected by multiple, frequency-selective channels that have a frequency bandwidth of approximately 1-2 octaves (Graham and Nachmias, 1971). A similar argument has been made for the existence of multiple independent shape channels that analyze complex outline shapes into sets of radial frequency components (Bell and Badcock, 2009; Bell *et al.*, 2007). To test this idea, two RF components can be combined along the same contour, and curvature detection thresholds for one RF component can be compared to detection thresholds measured when each RF component is displayed in isolation (Bell and Badcock, 2009; Bell *et al.*, 2007). Results from such pattern masking studies involving shapes demonstrate that the magnitude of interference observed is dependent upon the similarity in frequency between each RF component that together comprise the compound shape (Bell and

Badcock, 2009; Bell *et al.*, 2007). Lateral masking studies using RF contours show similar results, whereby maximal interference between a lateral mask and target RF contour is observed when the lateral mask is of similar frequency to the target contour (Habak *et al.*, 2004). Together, these results support an emergent view for the existence of independent shape channels that deconstruct complex forms into simpler constituent components.

Lateral masking has been used to estimate the size of perceptive fields, the psychophysical analog of receptive fields (Lev and Polat, 2015). Interestingly, perceptive fields and receptive fields often have similar properties: perceptive fields estimated in reverse-correlational studies often resemble receptive fields found in visual cortex (Lev and Polat, 2015; Neri and Levi, 2006). This result has led some researchers to infer the properties of receptive fields in visual cortex from perceptive field estimates derived from classical psychophysical techniques (Lev and Polat, 2015). Estimates of receptive field size in visual cortex using perceptive fields is made possible by assuming that the boundaries over which a receptive field is responsive can be estimated via the separation between two stimuli, namely a target and lateral mask (Lev and Polat, 2015). To estimate the receptive field size of V4 neurons, a visual area that is thought to be important for the perception of curvature, Habak *et al.* (2004) varied the separation between a target RF contour and lateral mask. As the separation between the lateral mask and target increased, the amount of lateral masking decreased (Habak *et al.*, 2004). From these results, Habak *et al.* (2004) estimated the perceptive field size for curvature-detecting mechanisms, which are presumably linked to V4 neurons, to be $\approx 1.14°$ in diameter. This value overestimates the receptive fields size of V4 neurons as measured via electrophysiological recording of V4 receptive field size

in macaques (Gattass *et al.*, 1988), but is a good approximation for a non-invasive behavioural method.

Spatial masking also serves as a useful tool for determining what features are important for the perception of shape of a target contour. For example, points of maximal curvature are thought to be of greater importance to the perception of objects relative to other shape features (Attneave, 1954; Barenholtz *et al.*, 2003; Biederman, 1987; Dickinson *et al.*, 2015, 2018; Loffler, 2008). Applied to RF contours, perception of the global form of a closed contour would be determined primarily by the shape's corners rather than its sides. In support of this proposition, Poirier and Wilson (2007) found that performance at detecting curvature along a target RF contour was most affected when select components of a lateral mask were made visible near the corners of target contours, with little masking being observed for lateral masks made visible parallel to the target's sides. Consistent with these observations, there is an increase in the strength of masking observed when the amplitude of a lateral RF mask also increases, despite the increasing misalignment between orientations of the elements that comprise sides of a target and mask (Habak *et al.*, 2004).

Temporal domain

More dynamic patterns of visual processing can be investigated with the addition of a temporal interval separating the presentation of a target and mask (Breitmeyer and Ogmen, 2000). Temporal masking paradigms differ from their spatial domain counterparts in the types of selective interference that are introduced along the visual pathway (Enns and Di Lollo, 2000). While spatial masking can be viewed as a problem centered around the summation of information across space, temporal masks

can result in both the inappropriate integration of features across space, but the additional capacity to completely disrupt the processing of visual information along the visual hierarchy (Breitmeyer and Ogmen, 2000; Enns and Di Lollo, 2000). Therefore, temporal masking paradigms serve as a useful tool for determining the time course required for the visual system to process and structure visual elements to form a percept (Enns and Di Lollo, 2000).

To determine the time course of interactions between shape representations within visual cortex, Habak *et al.* (2006) varied the stimulus onset asynchronies (SOA) that separated the presentation of a lateral mask and target RF contour. For both positive and negative SOAs, masking was observed across all SOAs suggesting a broad window of interaction between shape representations in visual cortex (Habak *et al.*, 2006). However, peak masking was observed when a mask was presented approximately 100 ms after the onset of a target shape (Habak *et al.*, 2006). After further investigation, Habak *et al.* (2006) were able to conclude this temporal window of interaction between shape representations may be a byproduct of the time course required for shape information to be processed by intermediate areas of visual cortex.

Although spatio-temporal interactions between shape representations have been previously examined (Bell and Badcock, 2009; Bell *et al.*, 2008, 2009; Habak *et al.*, 2004, 2006; Hess *et al.*, 1999; Jeffrey *et al.*, 2002; Kurki *et al.*, 2014; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998), there still remains a large gap in the literature on shape perception in understanding how the appearance of a shape is altered by the presence of a neighboring closed contour stimulus. To fill this void in the literature, the current dissertation examines what features of a laterally displaced shape play a critical role in determining the strength of interactions with a

target shape.

1.0.4   Thesis overview

In Chapter 2, we examine the the ability of the visual system to extract a global signal as curvature noise is introduced along a contour when more than one shape is present in a visual display. Objects seldom appear in isolation in natural scenes, and therefore it is important to understand how the mechanisms responsible for shape perception are able to handle noisey representations of form when more than one shape is being viewed. Previous work has shown that the ability to detect curvature along a target shape is made worse by the presence of a similar shape positioned lateral to the target shape (Habak *et al.*, 2004). Adopting a similar masking paradigm, observers were asked to detect deformations from circularity along a target contour while a second shape was positioned just lateral to the target shape. Over several experimental conditions, noise was introduced along the contour of the lateral mask, and the magnitude of noise required to alleviate masking effects were assessed. Results indicate that the visual system is robust at detecting low periodic curvature frequencies, but not as efficiently as when noise is added to isolated target shape (Bell *et al.*, 2007). These results also suggest that the addition of a high amplitude, high frequency curvature along a closed contour can modulate levels of shape interactions in multi-shape displays.

In Chapter 3, we examine whether the effect of phase alignment on the strength of masking changes as a function of the number of curvature maxima that occur per circumference of a circle. In addition, we examined whether patterns of phase dependent masking depend on sensitivity to phase shifts for single cycle contours, along with

what curvature features are of critical importance in modulating the strength of masking observed. For experimental procedures employing lateral masks, performance was assessed by estimating detection thresholds for curvature along the circumference of a target shape. Performance for sensitivity to shifts in phase for single cycle contours was assessed via thresholds for the minimum shift in phase needed for an observer to reliably discriminate the direction the shape was shifted (i.e., clockwise or counterclockwise). The effect of phase over the strength of masking observed did not differ as a function of the number of curvature maxima occurring along a closed contour. Differences in the effect of phase over patterns of phase dependent masking may be attributed to differences in sensitivity of observers at detecting shifts in relative phase of single cycles of modulation. Furthermore, positive curvature features contributed more to the full pattern of phase dependent masking observed compared to negative curvature extrema, although the presence of both curvature features are necessary to elicit maximal interactions between shapes. Together, these findings suggest that curvature frequency does not affect the strength of phase-dependent masking between shapes, but does influence observers' ability to discriminate phase shifts of single-cycle shapes.

In Chapter 4, we examined how sensitivity to deformations of curvature change as a function of polar angle, and to what extent angular frequency and spatial uncertainty might modulate such measures. The motivation for this study emerged from modelling efforts in Chapter 3, as previous computational models of the response properties of V4 neurons often assume that sensitivity to curvature remains constant across all polar angles (Pasupathy and Connor, 2001, 2002). Therefore, we tested this prediction using a novel stimulus call we call Difference of Gaussian contours which

allowed us to independently manipulate the amplitude, angular frequency, and polar angle of curvature of a closed-contour shape. Our results show that observers are, at first approximation, uniformly sensitive at detecting curvature deformations across all (24) polar angles tested. Furthermore, we found that this uniformity in sensitivity to curvature across polar angles remained robust to changes in angular frequency and spatial uncertainty. These results suggest that curvature deformations are equally well represented across all polar angles by neural mechanisms tuned to curvatures along the contour of closed shapes.

In Chapter 5, we examine how temporal masking modulates the spatial interactions that occur between two shapes. To accomplish this, a closed-contour temporal mask was presented at various stimulus onset asynchronies (SOAs) relative to the onset of a two-shape display consisting of a lateral mask and a target contour. Performance was assessed by examining how curvature detection thresholds in the target contour varied as a function of SOA. Results demonstrate that masking between the lateral mask and target contours was affected by the temporal mask only at intermediate, negative SOAs (i.e., -180 to -130 ms). Furthermore, the effects of temporal and spatial masks on curvature discrimination thresholds, at first approximation, are additive. Therefore, interactions between shapes that arise from temporal and spatial components can be described by a simple additive model that captures how each component affects perception of a target shape.

The General Discussion in Chapter 6 considers these results within the context of current findings in the field of shape perception, describes open research questions, and avenues for future directions of research. In summary, these studies show that sensitivity to curvature along closed contour shapes is affected by spatio-temporal

interactions that can occur between shape representation at different stages along the visual pathway.

1.1   References

Anderson, N. D. and Wilson, H. R. (2005). The nature of synthetic face adaptation. *Vision Research*, **45**(14), 1815–1828.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, **61**(3), 183–193.

Baldwin, A. S., Schmidtmann, G., Kingdom, F. A. A., and Hess, R. F. (2016). Rejecting probability summation for radial frequency patterns, not so Quick! *Vision Research*, **122**, 124–134.

Barenholtz, E., Cohen, E. H., Feldman, J., and Singh, M. (2003). Detection of change in shape: an advantage for concavities. *Cognition*, **89**(1), 1–9.

Bell, J. and Badcock, D. R. (2009). Narrow-band radial frequency shape channels revealed by sub-threshold summation. *Vision Research*, **49**(8), 843–50.

Bell, J., Badcock, D. R., Wilson, H., and Wilkinson, F. (2007). Detection of shape in radial frequency contours: independence of local and global form information. *Vision Research*, **47**(11), 1518–22.

Bell, J., Dickinson, J. E., and Badcock, D. R. (2008). Radial frequency adaptation suggests polar-based coding of local shape cues. *Vision Research*, **48**(21), 2293–2301.

Bell, J., Wilkinson, F., Wilson, H. R., Loffler, G., and Badcock, D. R. (2009). Radial frequency adaptation reveals interacting contour shape channels. *Vision Research*, **49**(18), 2306–17.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**(2), 115–147.

Bosking, W. H., Zhang, Y., Schofield, B., and Fitzpatrick, D. (1997). Orientation Selectivity and the Arrangement of Horizontal Connections in Tree Shrew Striate Cortex. *Journal of Neuroscience*, **17**(6), 2112–2127.

Bredfeldt, C. E. and Ringach, D. L. (2002). Dynamics of Spatial Frequency Tuning in Macaque V1. *Journal of Neuroscience*, **22**(5), 1976–1984.

Breitmeyer, B. G. and Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, **62**(8), 1572–1595.

Brincat, S. L. and Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience*, **7**(8), 880–886.

Carlson, E. T., Rasquinha, R. J., Zhang, K., and Connor, C. E. (2011). A Sparse Object Coding Scheme in Area V4. *Current Biology*, **21**(4), 288–293.

Chen, S. and Levi, D. M. (1996). Angle judgment: Is the whole the sum of its parts? *Vision Research*, **36**(12), 1721–1735.

Das, A. and Gilbert, C. D. (1995). Long-range horizontal connections and their role in cortical reorganization revealed by optical recording of cat primary visual cortex. *Nature; London*, **375**(6534), 780–4.

Dickinson, J. E., Bell, J., and Badcock, D. R. (2013). Near their thresholds for detection, shapes are discriminated by the angular separation of their corners. *PLoS ONE*, **8**(5), 1–9.

Dickinson, J. E., Cribb, S. J., Riddell, H., and Badcock, D. R. (2015). Tolerance for local and global differences in the integration of shape information. *Journal of Vision*, **15**(3), 21–21.

Dickinson, J. E., Haley, K., Bowden, V. K., and Badcock, D. R. (2018). Visual search reveals a critical component to shape. *Journal of Vision*, **18**(2), 2–2.

Downing, P. E., Jiang, Y., Shuman, M., and Kanwisher, N. (2001). A Cortical Area Selective for Visual Processing of the Human Body. *Science*, **293**(5539), 2470–2473.

Dumoulin, S. O. and Hess, R. F. (2007). Cortical specialization for concentric shape processing. *Vision Research*, **47**(12), 1608–1613.

Dumoulin, S. O. and Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage*, **39**(2), 647–660.

Elder, J. H. and Goldberg, R. M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision*, **2**(4), 5–5.

Enns, J. T. and Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, **4**(9), 345–352.

Field, D. J., Hayes, A., and Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research*, **33**(2), 173–93.

Foster, K. H., Gaska, J. P., Nagler, M., and Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *The Journal of Physiology*, **365**(1), 331–363.

Gallant, J. L., Jochen, B., and Van Essen, D. (1992). Selectivity for Polar, Hyperbolic, and Cartesian Gratings in Macaque Visual Cortex. *Neuroscience Letters*, **117**, 91.

Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W., and Essen, D. C. V. (1996). Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*, **76**(4), 2718–2739.

Gattass, R., Sousa, A. P., and Gross, C. G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *Journal of Neuroscience*, **8**(6), 1831–1845.

Geisler, W. S., Perry, J. S., Super, B. J., and Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, **41**(6), 711–724.

Gheorghiu, E. and Kingdom, F. A. A. (2006). Luminance-contrast properties of contour-shape processing revealed through the shape-frequency after-effect. *Vision Research*, **46**(21), 3603–3615.

Gheorghiu, E. and Kingdom, F. A. A. (2007). The spatial feature underlying the shape-frequency and shape-amplitude after-effects. *Vision Research*, **47**(6), 834–844.

Gilbert, C. D. (1977). Laminar differences in receptive field properties of cells in cat primary visual cortex. *The Journal of Physiology*, **268**(2), 391–421.

Gilbert, C. D. (1992). Horizontal integration and cortical dynamics. *Neuron*, **9**(1), 1–13.

Gilbert, C. D. and Wiesel, T. N. (1979). Morphology and intracortical projections

of functionally characterised neurones in the cat visual cortex. *Nature*, **280**(5718), 120–125.

Gilbert, C. D. and Wiesel, T. N. (1983). Clustered intrinsic connections in cat visual cortex. *Journal of Neuroscience*, **3**(5), 1116–1133.

Gilbert, C. D. and Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *The Journal of neuroscience*, **9**(7), 2432–2442.

Graham, N. and Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models. *Vision Research*, **11**(3), 251–IN4.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2017). Global processing of random-phase radial frequency patterns but not modulated lines. *Journal of Vision*, **17**(9), 18–18.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018a). Convergent evidence for global processing of shape. *Journal of Vision*, **18**(7), 7–7.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018b). The effect of spatiotemporal displacement on the integration of shape information. *Journal of Vision*, **18**(5), 4–4.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018c). Integration of shape information occurs around closed contours but not across them. *Journal of Vision*, **18**(5), 6–6.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., and Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, **24**(1), 187–203.

Grill-Spector, K., Knouf, N., and Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, **7**(5), 555–562.

Gustavsen, K. and Gallant, J. L. (2003). Shape perception: Complex contour representation in visual area V4. *Current Biology*, **13**(6), R234–R235.

Habak, C., Wilkinson, F., Zakher, B., and Wilson, H. R. (2004). Curvature population coding for complex shapes in human vision. *Vision Research*, **44**(24), 2815–23.

Habak, C., Wilkinson, F., and Wilson, H. R. (2006). Dynamics of shape interaction in human vision. *Vision Research*, **46**(26), 4305–20.

Heeley, D. W. and Buchanan-Smith, H. M. (1996). Mechanisms specialized for the perception of image geometry. *Vision Research*, **36**(22), 3607–3627.

Hegdé, J. and Essen, D. C. V. (2000). Selectivity for complex shapes in primate visual area V2. *Journal of Neuroscience*, **20**(5), RC61–RC61.

Hess, R. F., Wang, Y. Z., and Dakin, S. C. (1999). Are judgements of circularity local or global? *Vision Research*, **39**(26), 4354–60.

Hess, R. F., May, K. A., and Dumoulin, S. O. (2014). Contour integration: Psychophysical, neurophysiological, and computational perspectives. In J. Wagemans, editor, *The Oxford Handbook of Perceptual Organization*. Oxford University Press.

Hubel, D. H. and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, **148**(3), 574–591.

Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, **195**(1), 215–243.

Hubel, D. H., Wiesel, T. N., and Stryker, M. P. (1978). Anatomical demonstration of orientation columns in macaque monkey. *The Journal of Comparative Neurology*, **177**(3), 361–379.

Ito, M. and Komatsu, H. (2004). Representation of Angles Embedded within Contour Stimuli in Area V2 of Macaque Monkeys. *Journal of Neuroscience*, **24**(13), 3313–3324.

Jeffrey, B. G., Wang, Y. Z., and Birch, E. E. (2002). Circular contour frequency in shape discrimination. *Vision Research*, **42**(25), 2773–9.

Kempgens, C., Loffler, G., and Orbach, H. S. (2013). Set-size effects for sampled shapes: experiments and model. *Frontiers in Computational Neuroscience*, **7**.

Kennedy, G. J., Orbach, H. S., and Loffler, G. (2006). Effects of global shape on angle discrimination. *Vision Research*, **46**(8), 1530–1539.

Kennedy, G. J., Orbach, H. S., and Loffler, G. (2008). Global shape versus local feature: An angle illusion. *Vision Research*, **48**(11), 1281–1289.

Kobatake, E. and Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, **71**(3), 856–867.

Koffka, K. (1922). Perception: An introduction to the Gestalt-Theorie. *Psychological Bulletin*, **19**(10), 531–585.

Köhler, W. (1938). Physical Gestalten. In *A source book of Gestalt psychology*, pages 17–54. Kegan Paul, Trench, Trubner & Company, London, England. DOI: 10.1037/11496-003.

Kourtzi, Z. and Kanwisher, N. (2000). Cortical Regions Involved in Perceiving Object Shape. *Journal of Neuroscience*, **20**(9), 3310–3318.

Kourtzi, Z. and Kanwisher, N. (2001). Representation of Perceived Object Shape by the Human Lateral Occipital Complex. *Science*, **293**(5534), 1506–1509.

Kovács, I. and Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences*, **90**(16), 7495–7497.

Kurki, I., Saarinen, J., and Hyvarinen, A. (2014). Investigating shape perception by classification images. *Journal of Vision*, **14**(12), 1–19.

Lamme, V. A. F. and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, **23**(11), 571–579.

Lev, M. and Polat, U. (2015). Space and time in masking and crowding. *Journal of Vision*, **15**(13), 10–10.

Loffler, G. (2008). Perception of contours and shapes: Low and intermediate stage mechanisms. *Vision Research*, **48**(20), 2106–2127.

Loffler, G. (2015). Probing intermediate stages of shape processing. *Journal of Vision*, **15**(7), 1–1.

Loffler, G., Wilson, H. R., and Wilkinson, F. (2003). Local and global contributions to shape discrimination. *Vision Research*, **43**(5), 519–530.

Neri, P. and Levi, D. M. (2006). Receptive versus perceptive fields from the reverse-correlation viewpoint. *Vision Research*, **46**(16), 2465–2474.

Olshausen, B. A. and Field, D. J. (2005). How close are we to understanding V1? *Neural Computation*, **17**(8), 1665–1699.

Pasupathy, A. and Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, **82**(5), 2490–2502.

Pasupathy, A. and Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, **86**(5), 2505–2519.

Pasupathy, A. and Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience*, **5**(12), 1332–1338.

Poirier, F. J. and Wilson, H. R. (2006). A biologically plausible model of human radial frequency perception. *Vision Research*, **46**(15), 2443 – 2455.

Poirier, F. J. and Wilson, H. R. (2007). Object perception and masking: Contributions of sides and convexities. *Vision Research*, **47**(23), 3001–3011.

Prins, N., Kingdom, F. A. A., and Hayes, A. (2007). Detecting low shape-frequencies in smooth and jagged contours. *Vision Research*, **47**(18), 2390–2402.

Regan, D., Gray, R., and Hamstra, S. J. (1996). Evidence for a neural mechanism that encodes angles. *Vision Research*, **36**(2), 323–330.

Ringach, D. L., Shapley, R. M., and Hawken, M. J. (2002). Orientation Selectivity in Macaque V1: Diversity and Laminar Dependence. *Journal of Neuroscience*, **22**(13), 5639–5651.

Rockland, K. S. and Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *The Journal of Comparative Neurology*, **216**(3), 303–318.

Rust, N. C., Schwartz, O., Movshon, J. A., and Simoncelli, E. P. (2005). Spatiotemporal Elements of Macaque V1 Receptive Fields. *Neuron*, **46**(6), 945–956.

Schmidtmann, G. and Fruend, I. (2019). Radial frequency patterns describe a small and perceptually distinct subset of all possible planar shapes. *Vision Research*, **154**, 122–130.

Schmidtmann, G. and Kingdom, F. A. A. (2017). Nothing more than a pair of curvatures: A common mechanism for the detection of both radial and non-radial frequency patterns. *Vision Research*, **134**, 18–25.

Schmidtmann, G., Kennedy, G. J., Orbach, H. S., and Loffler, G. (2012). Non-linear global pooling in the discrimination of circular and non-circular shapes. *Vision Research*, **62**, 44–56.

Sincich, L. C. and Horton, J. C. (2005). The circuitry of V1 and V2: Integration of color, form, and motion. *Annual Review of Neuroscience*, **28**(1), 303–326.

Smith, A. T., Singh, K. D., Williams, A. L., and Greenlee, M. W. (2001). Estimating

Receptive Field Size from fMRI Data in Human Striate and Extrastriate Visual Cortex. *Cerebral Cortex*, **11**(12), 1182–1190.

Stettler, D. D., Das, A., Bennett, J., and Gilbert, C. D. (2002). Lateral Connectivity and Contextual Interactions in Macaque Primary Visual Cortex. *Neuron*, **36**(4), 739–750.

Ts'o, D. Y. and Gilbert, C. D. (1988). The organization of chromatic and spatial interactions in the primate striate cortex. *Journal of Neuroscience*, **8**(5), 1712–1727.

Ts'o, D. Y., Gilbert, C. D., and Wiesel, T. N. (1986). Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *Journal of Neuroscience*, **6**(4), 1160–1170.

Tyler, C. W. (1973). Periodic vernier acuity. *The Journal of Physiology*, **228**(3), 637–647.

Van Essen, D. C. and Gallant, J. L. (1994). Neural mechanisms of form and motion processing in the primate visual system. *Neuron*, **13**(1), 1–10.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., and von der Heydt, R. (2012a). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, **138**(6), 1172–1217.

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., and van Leeuwen, C. (2012b). A century of Gestalt psychology in visual

perception: II. Conceptual and theoretical foundations. *Psychological Bulletin*, **138**(6), 1218–1252.

Wertheimer, M. (1912). Experimentelle Studium uber das Sehen von Bewegung. *Zeitschrift fur Psychologie*, **61**(3), 161–265.

Wertheimer, M. (1923). Laws of Organization in Perceptual Forms. *Psycologische Forschung*, **4**, 301–350.

Westheimer, G. (1975). Editorial: Visual acuity and hyperacuity. *Investigative Ophthalmology & Visual Science*, **14**(8), 570–572.

Wilder, J., Fruend, I., and Elder, J. H. (2018). Frequency tuning of shape perception revealed by classification image analysis. *Journal of Vision*, **18**(8), 9–9.

Wilkinson, F., Wilson, H. R., and Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, **38**(22), 3555–68.

Wilkinson, F., James, T. W., Wilson, H. R., Gati, J. S., Menon, R. S., and Goodale, M. A. (2000). An fMRI study of the selective activation of human extrastriate form vision areas by radial and concentric gratings. *Current Biology*, **10**(22), 1455–1458.

Wilson, H. R. and Richards, W. A. (1989). Mechanisms of contour curvature discrimination. *Journal of the Optical Society of America A*, **6**(1), 106–115.

Wilson, H. R. and Richards, W. A. (1992). Curvature and separation discrimination at texture boundaries. *Journal of the Optical Society of America A*, **9**(10), 1653–1662.

Wilson, H. R. and Wilkinson, F. (2002). Symmetry perception: a novel approach for biological shapes. *Vision Research*, **42**(5), 589–97.

Wilson, H. R. and Wilkinson, F. (2015). From orientations to objects: Configural processing in the ventral stream. *Journal of Vision*, **15**(7), 4.

Wilson, H. R., Wilkinson, F., Lin, L. M., and Castillo, M. (2000). Perception of head orientation. *Vision Research*, **40**(5), 459–72.

Wilson, H. R., Loffler, G., and Wilkinson, F. (2002). Synthetic faces, face cubes, and the geometry of face space. *Vision Research*, **42**(27), 2909–2923.

Chapter 2

High periodic frequencies along a low radial frequency mask alleviates lateral masking effects

2.1   Abstract

Previous research suggests mechanisms that process curvature along shapes are able to detect low frequencies of curvature along a contour containing higher frequency components (Bell *et al.*, 2007). In this study, we examine whether shape mechanisms are still able to represent low frequency curvature in the presence of a higher frequency component along contours in multi-shape displays. Using radial frequency (RF) contours, we measured curvature detection thresholds along a RF5 target in the presence of a compound mask containing both a low (RF5) and high (RF25) RF component. In Experiment 1, we demonstrate that while masking is observed for conditions where only the RF5 component in the mask is made visible, no masking is observed for a

RF25 mask, and critically, also for a compound (RF5 + RF25) mask. Results from Experiment 2 show that masking effects progressively decline with increasing amplitude of the high frequency (RF25) component in the compound mask. Together, these results suggest that additions of high amplitude, high frequency curvature along a contour path can modulate the strength of interaction observed between shapes in multi-shape displays.

2.2   Introduction

The extraction of local and global contour curvature is fundamental for accurately perceiving shape (Bell *et al.*, 2007, 2009; Green *et al.*, 2017, 2018; Hess *et al.*, 1999; Jeffrey *et al.*, 2002; Loffler *et al.*, 2003; Wilder *et al.*, 2018; Wilkinson *et al.*, 1998). To better understand how the human visual system processes local contours to represent global shape, prior research has used Radial Frequency (RF) contours, which are visual stimuli generated by sinusoidally modulating the radius of a circle (Bell and Badcock, 2009; Bell *et al.*, 2009; Day and Loffler, 2009; Habak *et al.*, 2004, 2006, 2009; Hess *et al.*, 1999; Jeffrey *et al.*, 2002; Loffler *et al.*, 2003; Wilkinson *et al.*, 1998). RF contours are useful for studying shape perception because they can be varied parametrically to create shapes such as triangles and squares, as well as more complex objects such as faces (Wilson and Wilkinson, 2002; Wilson *et al.*, 2000). Human observers also have been shown to be hypersensitive at discriminating different RF contours (Birch *et al.*, 2000; Wilkinson *et al.*, 1998), suggesting that these stimuli are well represented in visual cortex (Habak *et al.*, 2004, 2009).

Evidence suggests that modulation of the curvature of circular contours is detected by independent local or global mechanisms (Green *et al.*, 2017, 2018; Hess *et al.*, 1999;

Jeffrey *et al.*, 2002; Loffler *et al.*, 2003), depending on the contour's radial frequency (i.e., the number of sinusoidal modulations per circumference, or 2π radians). Detection of radial frequencies between 2 and 8 cycles/2π relies on a global mechanism, as probability summation of the responses of local mechanisms fails to explain the threshold reduction that occurs with increasing numbers of cycles of radial modulation (Green *et al.*, 2017, 2018; Hess *et al.*, 1999; Jeffrey *et al.*, 2002; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). Kurki *et al.* (2014) have also shown using a classification imaging paradigm that contour parts of a low RF pattern (RF4) undergo linear, and not probability, summation. In contrast, a local mechanism may account for the detection of RF contours above 8 cycles/2π, as probability summation across local detectors does predict human observer performance (Hess *et al.*, 1999; Jeffrey *et al.*, 2002; Loffler *et al.*, 2003). Although, it still remains unclear how shape mechanisms integrate curvature segments to form a cohesive shape (see Baldwin *et al.* (2016)).

Masking paradigms have been used to examine the interaction among mechanisms that detect different RF contours. Previous studies have shown that very little masking is observed between low and high RF contours, which is consistent with the idea that radial frequencies are detected by multiple, RF-tuned mechanisms (Bell *et al.*, 2007, 2009; Habak *et al.*, 2004). For example, Bell *et al.* (2007) have shown that detection threshold for an RF3 contour in the presence of a coincident, supra-threshold RF24 contour on the same contour (i.e., an RF3 + RF24 compound contour), is the same as the detection threshold for an RF3 contour by itself. These results suggest that shape-encoding mechanisms that represent low-frequency curvature, and detect RF3 patterns, are able to tolerate high frequency perturbations in local curvature (e.g., orientation and position cues). Within the context of our discussion, masking

will be treated as a proxy for measuring the strength of interactions between shapes, and herein the two will be used interchangeably. Thus, a consequence of low and high RF components being represented independently is the lack of interaction (i.e., interference) between high and low curvature components along the same (Bell *et al.*, 2007, 2009) or separate (Habak *et al.*, 2004) contours. However, the frequency at which high and low frequency curvatures no longer interact remains unknown, as masking is still observed between RF3 and RF11 patterns coincident along the same contour (Bell *et al.*, 2009).

Most studies of RF masking have used compound RF contours: the mask and target RF components were on a single contour. However, most natural scenes contain multiple objects and contours, and therefore it is important to understand how perceptual representations of objects and curvature may change when more than one object outline appears in an observer's field of view. One approach taken in examining how shape and curvature perception is potentially altered in multi-shape displays is to examine how RF interactions occur between contours. Psychophysical studies have shown that interactions between shapes can lead to higher curvature detection thresholds (Bell *et al.*, 2007; Habak *et al.*, 2004, 2006, 2009), but little is known about how the RF specificity of masking is affected. The current study investigated how sensitivity to low RF curvature was affected by low and high RF masking components in multi-shape displays. Experiment 1 measured curvature detection thresholds for an RF5 contour that appeared alone, and in the presence of RF mask components that were added to a separate contour (see Figure 2.1). The mask contour contained a low radial frequency (RF5) that was set to match the target frequency (RF5), and/or a high radial frequency (RF25) that perturbed the low-frequency curvature

along the masking contour. If between-contour masking is the same as masking along a single contour, then we expected to find that detection thresholds for RF5 target curvature would be increased by a lateral mask consisting of RF5, but not RF25, curvature. Critically, from these predictions, a compound (RF5 + RF25) mask would also produce the same amount of masking as RF5 alone.

2.3 Methods

2.3.1 Participants

The final sample of participants comprised seven experienced psychophysical observers ($M = 23.71$ years; $SD = 2.36$) across two experiments. Five of the participants were naïve to the goals of the study (TUI, BTP, MCT, AAR, OLI), one participant was aware of the purpose behind the study (MP), and one was an author (MS). All participants had normal or corrected-to-normal visual acuity. Experimental protocols were approved by the McMaster University Research Ethics Board, and informed consent of the participant was collected prior to the start of the experiment.

2.3.2 Apparatus & Stimuli

An APPLE G4 2.66 GHz Quad-Core Intel Xeon computer generated the stimuli and controlled the experiment using Matlab 10.7.0 (Mathworks Inc., Natick, MA) and Psychophysics and Video Toolbox (Brainard, 1997; Pelli, 1997). The visual display was a Sony Model GDM-F520 graphics monitor, with a pixel resolution 1024 x 768 (62 pixels/deg) and a refresh rate of 100 Hz. The display had a mean luminance of 65.6 cd/m^2, and was the only light source in the room. Stimuli were viewed binocularly

from a viewing distance of 131 cm, which was maintained through the use of a chinrest.

RF contours used in the current experiment are shown in Figure 2.1. Contours were similar to those used in previous research (Bell and Badcock., 2009; Bell et al., 2009, 2007; see Wilkinson et al., 1998 for original) and were defined by the equation:

$$r(\theta) = \bar{r}(1 + A_1 sin(\omega_1\theta + \phi_1) + A_2 sin(\omega_2\theta + \phi_2)) \tag{2.1}$$

where θ is the polar angle, $\bar{r}$ is the mean radius of the pattern, and the modulation amplitude, radial frequency, and phase of the two RF components are represented by A, ω, and ϕ, respectively. To avoid having a contour cross the center of the pattern, the modulation amplitude of each RF component never exceeded 0.5. The radial luminance profile of each contour was defined by a fourth derivative Gaussian (D4: see Wilkinson *et al.* (1998)) with a luminance contrast set to 99% and a peak spatial frequency of 8 cpd. Both target and comparison contours had mean radius $\bar{r} = 1.14°$.

Stimuli were composed of two RF components, $\omega_1 = 5$ (RF5) and $\omega_2 = 25$ (RF25), and the amplitudes A_1 and A_2 were adjusted separately to create RF5, RF25, and compound (RF5 + RF25) masking contours. Target stimuli were RF5 contours (i.e., the amplitude of the RF25 component was set to zero). Setting both amplitudes to zero resulted in a circular (RF0) contour. Orientation of targets was varied across three relative target-mask phase combinations (0°, 90°, 180°). For baseline measurements, where no mask appeared, targets were shown at three orientations (45°, 135°, 225°) that differed from those used in the masking conditions.

For masks containing a single RF component (RF5 or RF25), RF amplitude was set at 15 times detection threshold. For compound masks, the amplitudes of both RF components were set to 15 times detection threshold. An unmodulated (RF0) contour
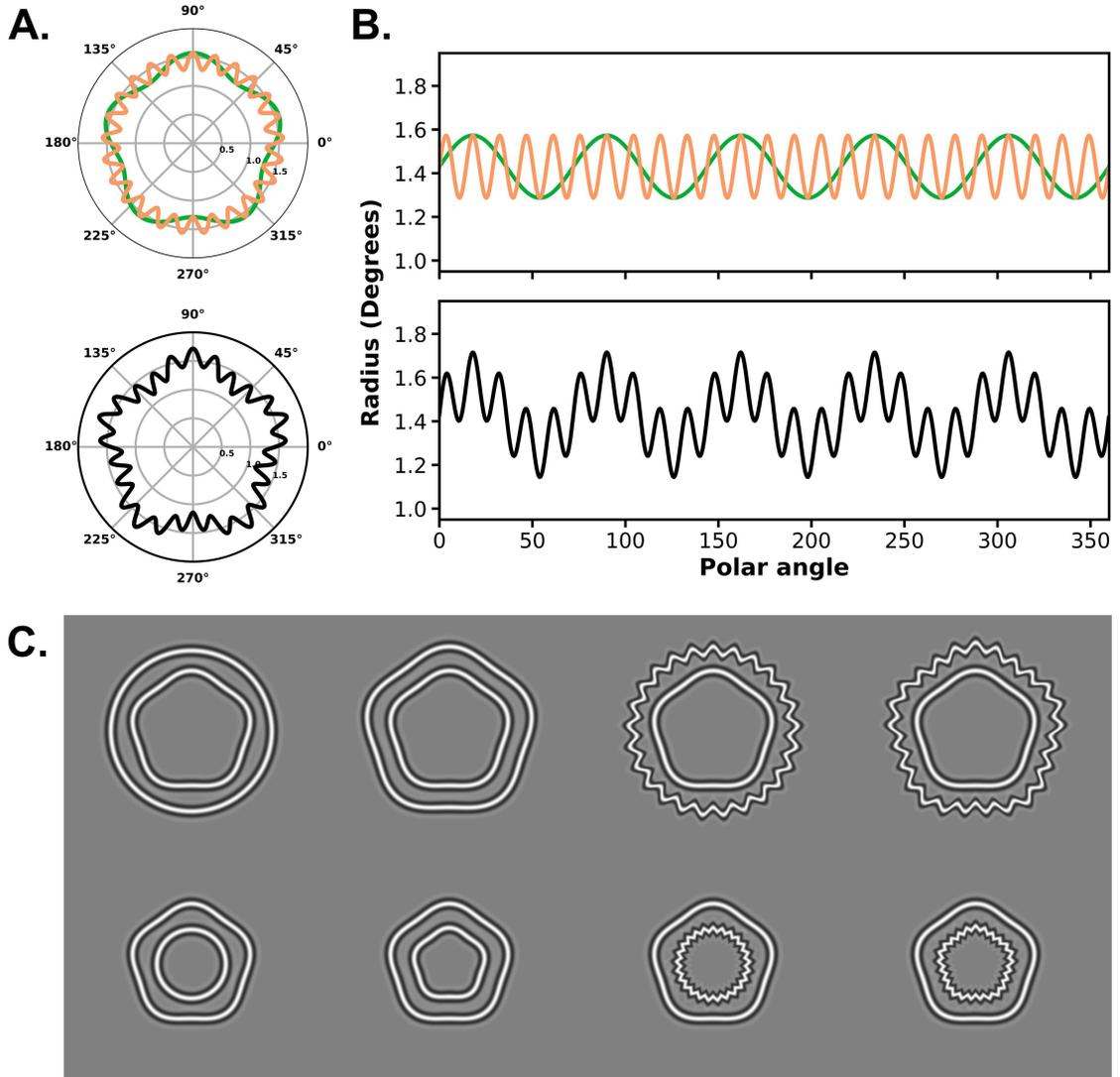
Figure 2.1: (A) Polar plots of modulated RF masks used in Experiment 1. *Top:* RF5 (green) and RF25 (orange) masks. *Bottom:* Compound RF mask (black) composed of RF5 and RF25 components. (B) Plot of radius (visual degrees) of modulated masks as a function of polar angle (degrees). Colours correspond to same RF contours as in Figure 1A. (C) Example of the four types of RF masks used positioned outside of a target RF5 contour. Masks from left to right: Control, RF5, RF25, and RF5+25 contour. RF contours are shown at a relative target-mask phase of 0°, with contrast at 99%. For illustrative purposes, all RF components are shown are at a 10% modulation amplitude.

served as a control mask. Masks were positioned either inside or surrounding (outside) the target contour, with an average distance of 0.29° maintained across contours within and between conditions. In the control condition with an unmodulated mask (RF0), the target-mask distance was set to equal the minimum distance that appeared across masking conditions between contours. Across all conditions, the phase of the mask was held constant.

2.3.3   Psychophysical Procedure

The experiment used a two-interval forced choice paradigm: one interval contained a target contour with RF5 curvature and the other interval contained a circular (RF0) target contour. The order of the intervals was randomized across trials, and observers indicated which of the two intervals contained the RF5 target contour by pressing one of two keys on a computer keyboard. Each trial was initiated by an observer by pressing the spacebar key. Upon initiating a trial, a fixation dot flickered in the center of the screen for 250 ms followed by a 200 ms delay. Each interval was presented onscreen for 150 ms with a 300 ms ISI. The spatial position of each stimulus was jittered on each interval within a 0.17° radius from the center of the screen. Prior to the start of a block of experimental trials, observers received 10 practice trials with auditory feedback to ensure they understood the task and were familiar with the stimuli. Practice trials contained the same set of stimuli that would appear in the experimental trials except that target stimuli were shown at greater modulation amplitudes than those of the main experiment. A 60 s light adaptation period followed the completion of practice trials, after which the experiment started.

The target's modulation amplitude was varied across trials using the method of

constant stimuli, where each threshold reflects the minimum deformation amplitude required to differentiate between a modulated and unmodulated RF contour. A single session was comprised of six different modulation amplitudes, each shown 36 times at random, for a total of 216 trials. The presentation of different target-mask phase combinations was randomly interleaved within a session. Observers completed a minimum of two sessions for each mask condition and location.

2.3.4   Data analysis

All analyses reported within this paper were performed with R (R Core Team, 2017). Data from each experimental session were fit with a psychometric function defined by:

$$\psi(x; \alpha; \beta; \gamma; \lambda) = \gamma + (1 - \gamma - \lambda)F_W(x; \alpha, \beta) \tag{2.2}$$

where λ defines the lapse rate of an observer, γ the guess rate, and $F_W(x; \alpha, \beta)$ is a Weibull function (Weibull, 1951) described by:

$$F_W = 1 - exp(-(x/\alpha)^\beta) \tag{2.3}$$

Thresholds were defined as the modulation amplitude that yielded 75% detection accuracy. The strength of masking was defined as $\log_e(t_m/t_0)$, where t_m and t_0 are thresholds in the masked and unmasked conditions, respectively. All log-ratios (and thresholds) passed tests for normality, homogeneity of variance, and sphericity prior to fitting linear models.

2.4   Results

Masking strength is plotted as a function of masking condition in Figure 2.2 Positive values indicate detection thresholds for a given condition are elevated relative to baseline; negative values indicate thresholds are lower compared to baseline. Across all observers, thresholds obtained with an RF0 mask (i.e., a circular mask with no RF modulation) were only slightly higher than thresholds in the absence of a mask. Relative to the RF0 mask condition, four of the five observers had more masking with an RF5 mask than with either an RF25 mask or a compound (RF5 + RF25) mask. The fifth observer (TUI) had similar amounts of masking across most conditions. Across all observers and mask conditions, there was no obvious or systematic difference between masking strength obtained with masks located inside or outside of the target contour.

These observations were evaluated quantitatively with a 4 (Mask Type) × 2 (Mask Location) × 2 (Test Session) repeated-measures ANOVA performed on log-ratio data. The ANOVA revealed a significant main effect of Mask ($F_{3,12} = 12.29$, $p < .001$, $\eta_p^2 = .75$) and Session ($F_{1,4} = 31.23$, $p < .01$, $\eta_p^2 = .89$), with no effect of Mask Location ($F_{1,4} = 1.26$, $p = .32$, $\eta_p^2 = .24$). None of the interactions were significant ($F \leq 2.77$, $p \geq 0.087$, in all cases).

Post-hoc comparisons between masking conditions were performed using paired, two-sided t tests with p values adjusted with the Holm-Bonferonni method and familywise $\alpha = 0.05$ (Holm, 1979). The comparisons revealed significant differences between detection thresholds in the RF5 mask condition and remaining three mask conditions ($p_{Holm} < .001$), and between the RF25 and compound mask conditions ($p_{Holm} = 0.012$). All remaining comparisons were not significant.
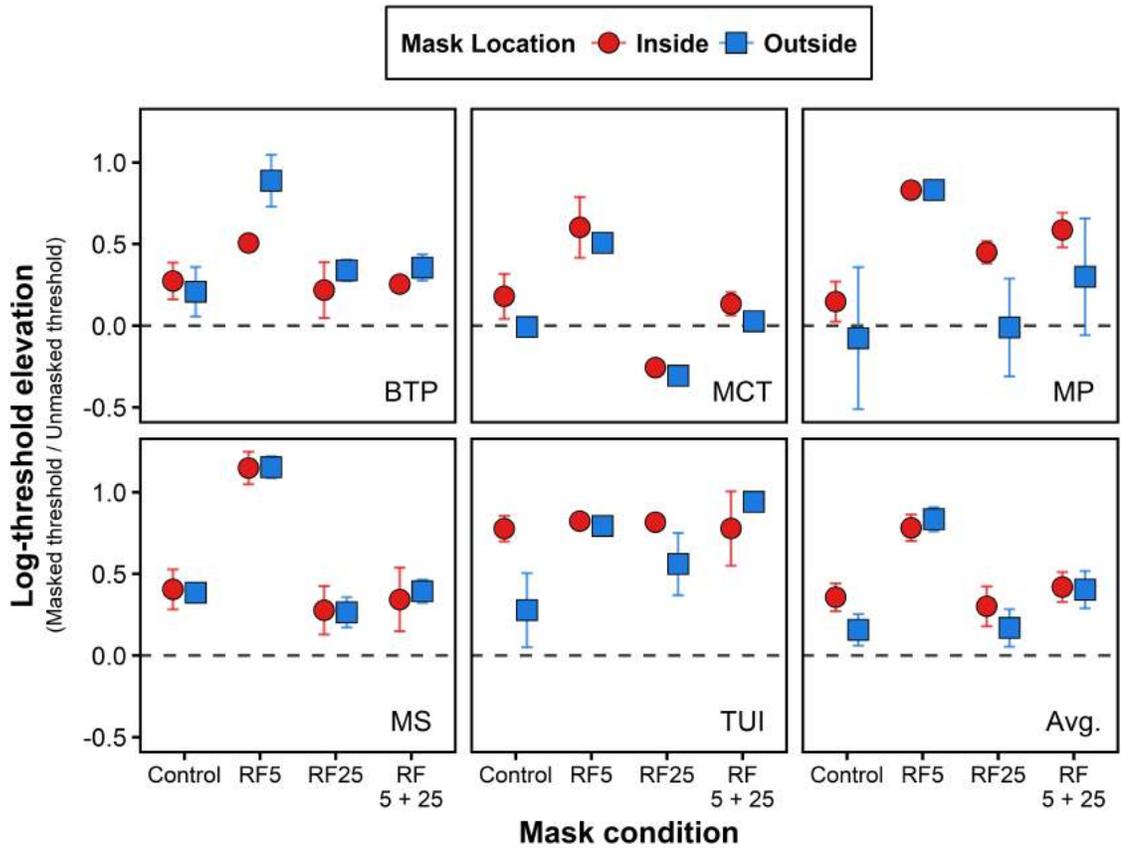
Figure 2.2: Masking, defined as the logarithm of the ratio of masked and unmasked detection thresholds, are plotted for each condition and observer. Baseline measures correspond to a log-ratio of zero, as indicated by horizontal dashed lines. Thresholds obtained with masks located inside or outside of the target contour are indicated by red and blue symbols, respectively. Error bars indicate ± 1 SEM.

To further elucidate the relation between masking strength and the modulation amplitude of the high RF component (RF25) in the compound mask, a second experiment measured RF5 detection thresholds with compound masks in which the amplitude of the RF25 component varied across five levels (i.e., 0.5, 1, 5, 15, and 25 times RF25 detection threshold). Examples of RF contours used in the second experiment are shown in Figure 2.3. In these conditions, mean target-mask distance was set to 0.40° to ensure that targets and masks did not overlap in any condition.

Figure 2.4 displays results from Experiment 2 as log-ratios between detection thresholds for masked and unmasked conditions as a function of RF25 modulation amplitude. Across all observers, masking strength decreased as RF25 modulation amplitude increased. For two observers (OLI and MS), detection thresholds at higher RF25 modulation amplitudes were comparable to thresholds measured without a mask.

The masking data from Experiment 2 were analyzed with linear mixed models that controlled for the fact that observers MS and AAR participated in one more test session per condition than the other observers. Two fixed effects (RF25 amplitude and Mask location) and two random effects (Observer and Session) were modelled. An ANOVA with Kenward-Roger (Kenward and Roger, 1997) approximations for degrees of freedom was used, as this method rescales the F ratios in addition to adjusting the degrees of freedom to better approximate F distributions (Judd *et al.*, 2012). The ANOVA revealed a significant main effect of RF25 amplitude ($F_{4,81} = 20.60$, $p < .0001$); the effect of Mask location ($F_{1,81} = 1.01$, $p = .32$) and the interaction between RF25 amplitude and Mask location ($F_{4,81} = 1.06$, $p = .38$) were not significant.

A.

B.
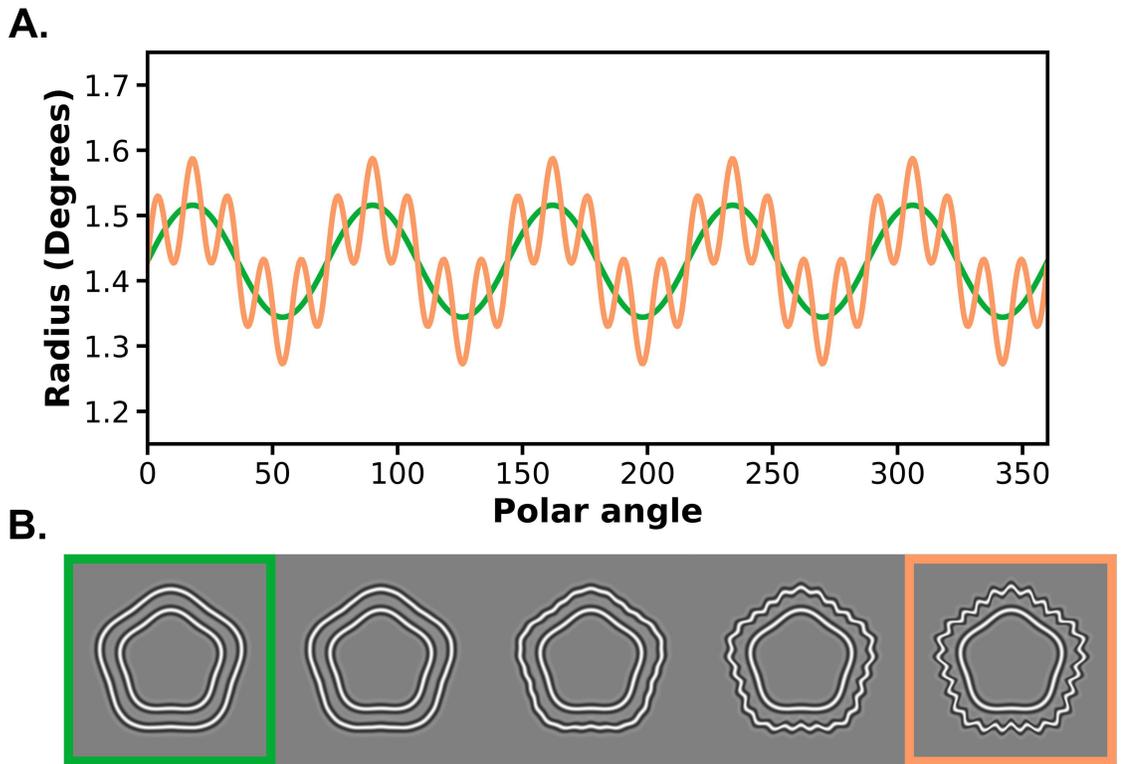
Figure 2.3: Illustration of compound RF5+RF25 masks with different RF25 modulation amplitudes used in second experiment. (A) Plot of minimum (green) and maximum (orange) RF25 amplitude modulations used in the experiment. (B) Luminance profile of compound RF masks: from left to right, the modulation amplitude of the RF25 component was set to 0.5, 1, 5, 15 ,and 25 times the RF detection threshold for observer MCT. The modulation amplitude of the RF5 component was always set at 15 times MCT's RF5 detection threshold.
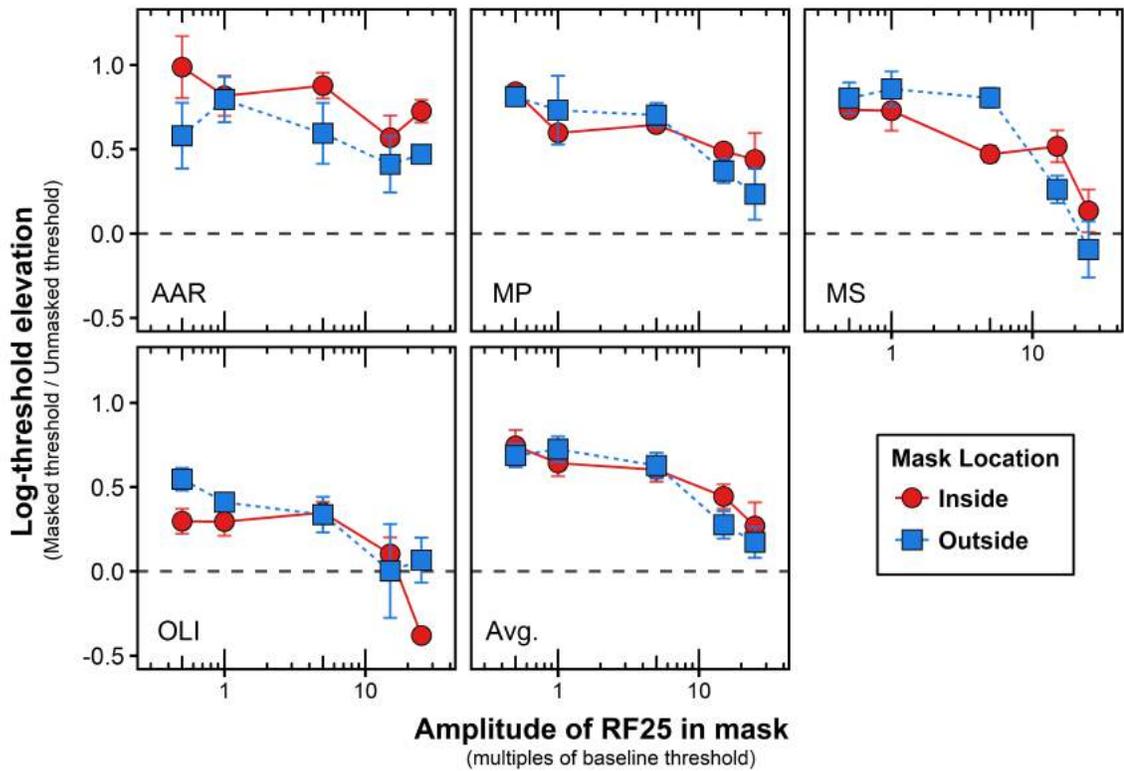
Figure 2.4: Log-ratio of RF5 detection thresholds between mask and unmasked conditions plotted as a function of RF25 modulation amplitude in a compound mask (RF5+RF25). Log-ratios computed for masks located inside or outside of the target contour are indicated by red and blue symbols, respectively. Error bars indicate ± 1 SEM. Baseline thresholds obtained for isolated RF5 contours are represented by horizontal dashed lines, and correspond to a log-ratio of zero.

Post-hoc comparisons between conditions were performed using the same Holm-Bonferonni procedure as used in Experiment 1, where familywise $\alpha = 0.05$ (Holm, 1979). Comparisons revealed significant differences between masking obtained when the RF25 component was 15 or 25 times threshold compared to all remaining conditions (all p_{Holm} values < 0.01). No other comparisons were significant ($p \geq 0.18$).

Experiment 2 examined whether introducing progressively higher amplitude, high frequency curvature along a contour affects the strength of interaction (i.e., strength of masking) between shapes. Our results demonstrate that at low modulation amplitudes, the RF25 component of the compound mask does not interfere with the ability of shape mechanisms at representing low frequencies of curvature, as masking is observed. As modulation amplitude of the RF25 component increases, the RF5+RF25 compound produces progressively less masking of the RF5 target.

2.5   Discussion

The aim of the current study was to determine the robustness of shape mechanisms to introductions of high frequency perturbations along the contour of a shape, in this case a visual mask, in multi-shape displays. Experiment 1 demonstrated that an RF5 contour has large masking effects on detection thresholds for a spatially separate RF5 target. In contrast, an RF25 contour, and surprisingly, also an RF5+RF25 compound contour has a very small masking effect on detection thresholds of an RF5 target. Furthermore, results from Experiment 2 show that introducing low amplitude, high frequency curvature along the contour of a mask does not affect strength of masking between shapes. However, at higher RF25 amplitudes, no interaction between shapes is observed. Our results suggest that additions of high amplitude, high frequency

curvature along contours affects interactions between low frequency shapes in multi-shape displays.

2.5.1 Visibility of RF5 in compound mask

One possible explanation of the results from Experiment 2 is that the RF25 component reduces visibility of the RF5 component in the compound mask. Introducing high frequency curvature along a mask can alter important curvature features that are used by the visual system to discriminate between different shapes. For example, local contour features in the compound mask, including measures of local orientation and position, are increasingly altered at higher RF25 amplitudes. Previous studies have found that both local contour orientation and position cues contribute to representations of concentric shapes (Day and Loffler, 2009; Dickinson *et al.*, 2013; Hess *et al.*, 1999; Wang and Hess, 2005) and extended contours (Field *et al.*, 1993; Hess *et al.*, 2003). However, if the presence of high amplitude, high RF curvature altered orientation and position cues sufficiently, then high RF curvature ought to reduce the visibility of low RF curvature on the same contour. Contrary to this prediction, data from Bell *et al.* (2007) argues against this idea, as very little masking has been found between low and high RF components along a single contour. Furthermore, in the current study, the amplitude of both RF components were at suprathreshold levels when the effect of the mask was diminished. Therefore, it is unlikely that altering local contour orientation and position by adding high frequency (RF25) curvature along a compound (RF5 + RF25) mask reduces visibility of the RF5 component.

In addition to altering local contour curvature, introducing high frequency curvature along a compound mask can modify important global properties of shape that

are used in detecting separate RF components. For example, as illustrated in Figure 2.5, angular separation between convex maxima from the origin of a shape can be a useful feature in representing different curvature frequencies (Dickinson *et al.*, 2013, 2015, 2018; Poirier and Wilson, 2006). Data from Dickinson *et al.* (2013) demonstrate that human observers are unable to accurately discriminate shapes if the periodicity of curvature maximum across shapes is similar. Therefore, perhaps at higher amplitudes, the RF25 component somehow makes it more difficult to make use of periodicity of curvature maximum as a cue to shape. Periodicity of curvature maximum can also be altered depending upon the phase relationship between RF components if coincident along a contour. However, an examination of Figure 2.6 suggests that disruptions in angular positions of curvature extrema do not occur for our stimuli. Here, we represent curvature as the the change in orientation of a contour relative to a circle defined by the second derivative of the angular function:

$$r(\theta)'' = -A_1\omega_1^2 sin(\omega_1\theta) - A_2\omega_2^2 sin(\omega_2\theta) \qquad (2.4)$$

Because the two RF components of the mask are harmonically related, and phase aligned, the position of global curvature extrema occur at the same polar angles as they would appear if the higher frequency component were not added. Data from Bell *et al.* (2007) also argue against the importance of periodicity of curvature maximum in detecting RF components along a compound shape, as detection thresholds for an RF3 pattern are unaffected by displacements in the angular position of curvature maxima by the addition of a phase randomized, RF24 component along the same contour (Bell *et al.*, 2007). Therefore, we think it is unlikely that the reduction in masking that occurs with increasing RF25 amplitude is caused by the RF25 component masking
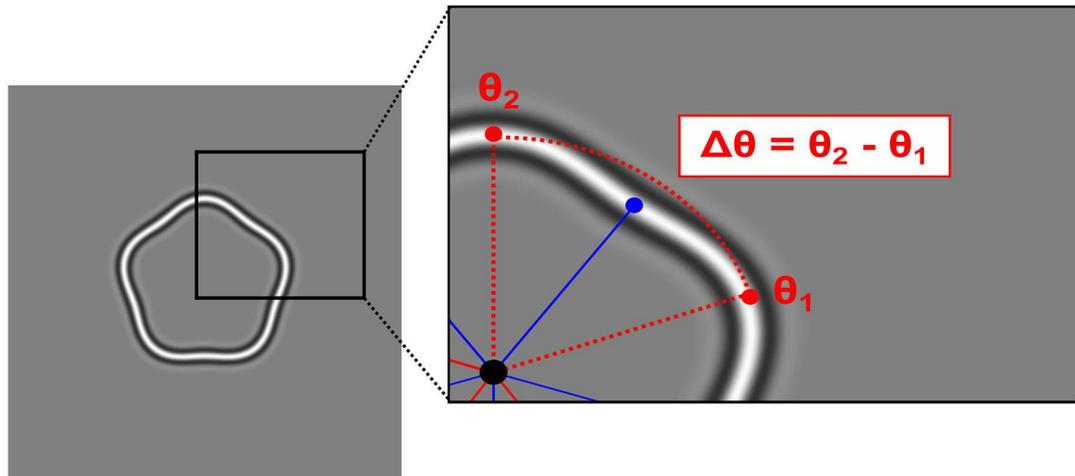
Figure 2.5: Example of a RF5 contour being represented by the angular separation between curvature maximum from the origin of the shape, as illustrated by the red semi-circle. Angular separations between cuvature minima, as illustrated in blue, may also be a useful feature in discriminating shapes.
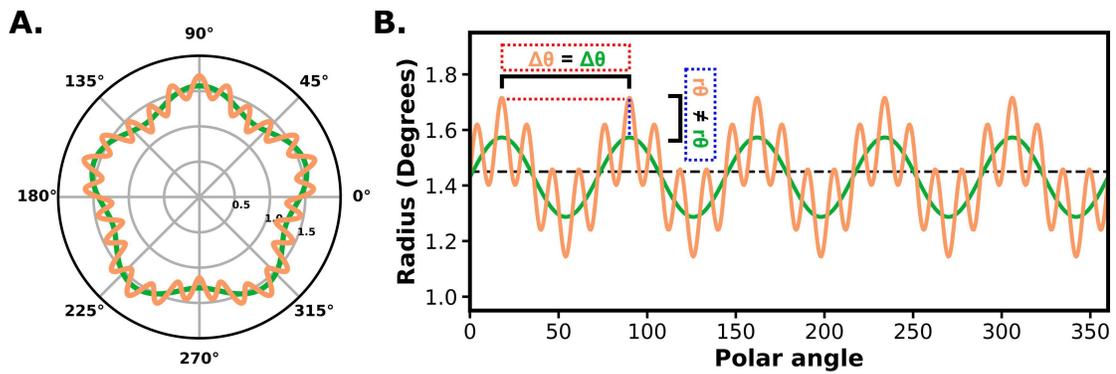
Figure 2.6: (A) Polar plot of RF5 (orange) and RF5+RF25 (green) masks used in Experiments 1 and 2. (B) Radius of masks in visual degrees plotted as a function of polar angle. These figures illustrate that important features of shape, such as angular separation between curvature extrema, remain unchanged with the introduction of a high frequency component (RF25) along the contour of a RF5 mask. However, radial distance of curvature extrema change with additions of higher frequency of curvature along a mask.

the RF5 component by disrupting either local or global shape properties. Instead, introducing high frequency curvature along a mask is likely altering other factors that contribute to levels of interaction between shapes, such as the radial distance between curvature maxima (Habak *et al.*, 2004, 2006; Poirier and Wilson, 2006) and alignment of sides (i.e., regions between curvature extrema) (Hess *et al.*, 1999; Poirier and Wilson, 2007).

2.5.2   Radial distance between curvature maxima

Another possibility is that the decline in masking illustrated in Figure 2.4 can be explained by an increase in the radial distance between curvature maxima of the target contour and compound mask as the amplitude of the RF25 component increases. Habak *et al.* (2004) demonstrated using spatially separated masks that strength of masking declines as mean target-mask distance is increased. Therefore, increasing distance between shapes also serves to decrease the level of interaction observed between these contours. While mean target-mask distance remains fixed across all of our experimental conditions, the distance between curvature extrema across shapes varied across conditions. As seen in Figure 2.7, increasing the amplitude of the RF25 component in the compound mask displaces curvature features both toward and away from the target contour. At sufficient separations in radial distance between positive curvature extrema, masking is no longer observed, as demonstrated in Experiments 1 and 2. This occurs despite the fact that the distances between mask and target curvature minima decrease at higher RF25 amplitudes.

Our results lend support for models of curvature-encoding V4 neurons that respond to concentric forms primarily by pooling positive changes in orientation across
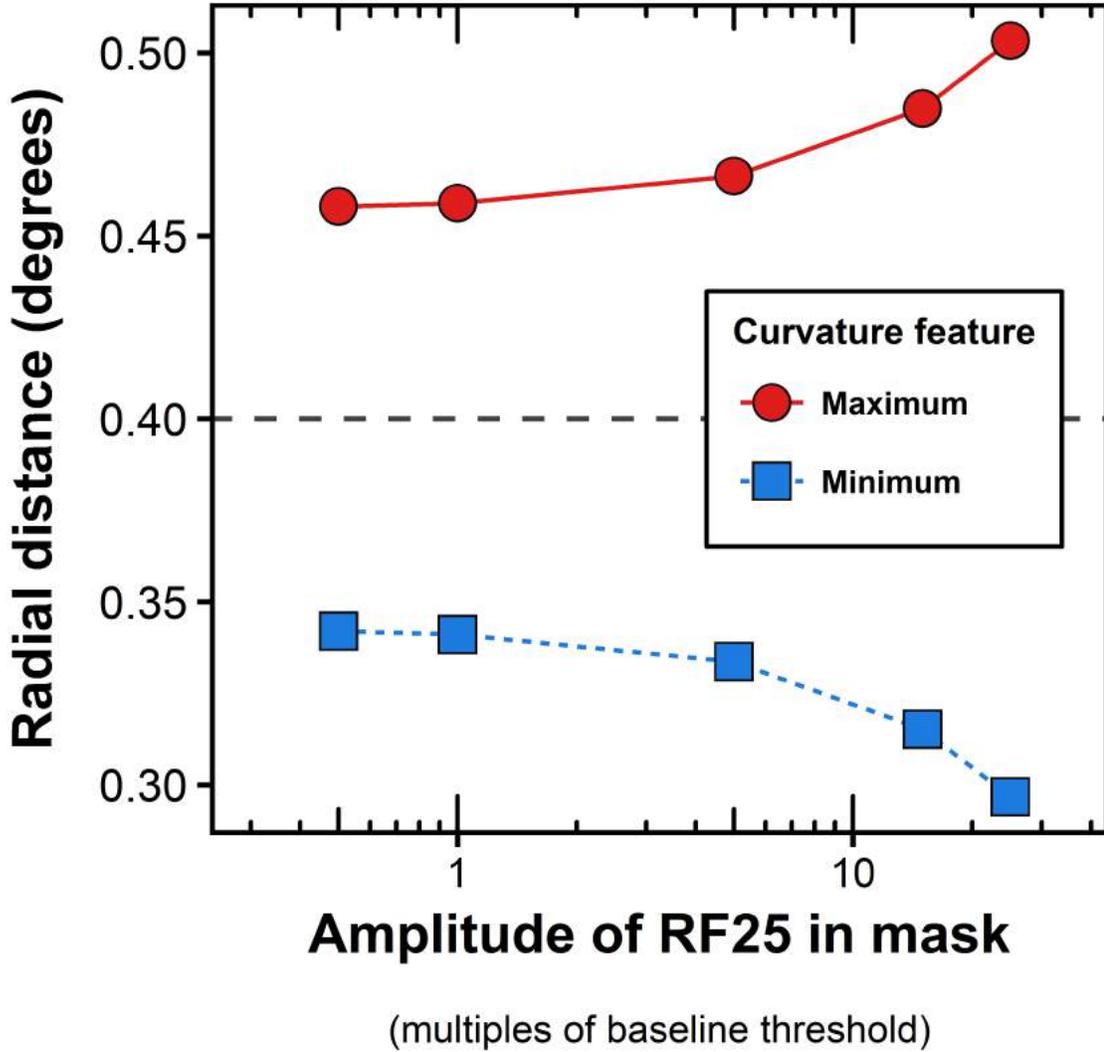
Figure 2.7: Plot of radial distance (in visual degrees) for maxima and minima curvature features between a RF5+RF25 mask and target RF5 shape as amplitude of the RF25 component in the mask varies. Mean baseline thresholds (i.e., no mask) across observers in Experiment 2 were used to compute the amplitude, and thus radial distances, for each RF component (RF5 and RF25). For curvature features, radial distance between positive maxima of a target and compound mask grow lager with increases in amplitude of the RF25 component in the mask. The opposite relation holds true of radial distance and negative minima.

their receptive fields (Habak *et al.*, 2004, 2009; Poirier and Wilson, 2006). If the response function of such units to curvature is a compressive non-linearity, then each subsequent curvature feature will need to be of higher strength (amplitude) to elicit a similar change in activation (Habak *et al.*, 2004, 2009). Thus, such models predict that larger modulations of curvature are required for an observer to correctly detect curvature along a target contour surrounded by a shape of similar form. As more convex features fall within overlapping V4 receptive fields, the response will saturate and a much larger curvature signal will be required to evoke a change in the response (Habak *et al.*, 2004). However, as the distance between positive curvature extrema increase, fewer curvature features will fall within overlapping receptive fields, and therefore sensitivity to curvature will be restored. Consequently, less curvature will be required for an observer to correctly detect deformations along the target contour (Habak *et al.*, 2004). Our results are consistent with these predictions.

2.5.3   Sides

Sides, here defined as the regions between curvature extrema (i.e., zero-crossings and otherwise), have also been shown to contribute to the effect of masking between multiple shapes (Hess *et al.*, 1999; Poirier and Wilson, 2007). Masking from sides is hypothesized to be caused by weak local inhibition occurring between orientation-selective filters that reduce neural responses for parallel contour segments (Poirier and Wilson, 2006, 2007). Furthermore, straight line segments are more susceptible to suppression than curved contour segments (Li *et al.*, 2000). This hypothesis is consistent with data from Experiments 1 and 2 that demonstrate that masking is significantly reduced when local orientation of the mask becomes increasingly less parallel to the

target at higher levels of perturbation. However, it is unlikely that interference due to the inhibitory interaction between parallel line segments is responsible for the pattern of masking observed across both sets of experiments. First, a perfectly circular mask contains orientation information more similar to a target shape at threshold levels of modulation relative to fully modulated masks, yet does not elevate detection thresholds (Habak *et al.*, 2004). By decomposing masks into parts that contain either convexities, sides, or both, Poirier and Wilson (2007) showed that interference caused by masks composed entirely of sides is less than that of full masks, or those composed of only convex parts. Together, these results suggest that masking observed in Experiments 1 and 2 are only partially the result of interference from sides, and thus interactions occurring between curvature features likely originates at higher levels of shape processing.

2.6 Conclusion

The aim of the current study was to examine how sensitivity to low RF curvature was affected by low and high RF masking components in multi-shape displays. Results from Experiment 1 demonstrate that masking is observed between a RF5 target and RF5 mask, but not for a RF25 mask, and surprisingly neither for a compound (RF5 + RF25) mask. Data from Experiment 2 show that when the high frequency (RF25) component in a compound mask (RF5+RF25) appears at low amplitudes, masking is observed between shapes. However, as amplitude of the high frequency component progressively increases, the compound mask is rendered ineffective. While introducing high frequency curvature along a mask alters local and global properties of curvature and shape, it is unlikely that these factors are responsible for the lack

of interaction observed between shapes at higher RF25 amplitudes. Instead, our data further support the idea that masking strength between shapes depends upon the radial distance between curvature maxima (Habak *et al.*, 2004, 2009; Poirier and Wilson, 2006), and that the alignment of sides may also be partly responsible in modulating masking strength between shapes (Hess *et al.*, 1999; Poirier and Wilson, 2007). Together, our results suggest that low and high frequency curvatures can interact in modulating the strength of interaction observed between two low curvature shapes. Future work should aim to examine the effect of phase at modulating the strength of interaction between shapes defined by multiple curvature frequencies.

2.7   References

Baldwin, A. S., Schmidtmann, G., Kingdom, F. A. A., and Hess, R. F. (2016). Rejecting probability summation for radial frequency patterns, not so Quick! *Vision Research*, **122**, 124–134.

Bell, J. and Badcock, D. R. (2009). Narrow-band radial frequency shape channels revealed by sub-threshold summation. *Vision Research*, **49**(8), 843–50.

Bell, J., Badcock, D. R., Wilson, H., and Wilkinson, F. (2007). Detection of shape in radial frequency contours: independence of local and global form information. *Vision Research*, **47**(11), 1518–22.

Bell, J., Wilkinson, F., Wilson, H. R., Loffler, G., and Badcock, D. R. (2009). Radial frequency adaptation reveals interacting contour shape channels. *Vision Research*, **49**(18), 2306–17.

Birch, E. E., Swanson, W. H., and Wang, Y. Z. (2000). Infant hyperacuity for radial deformation. *IOVS*, **41**(11), 3410–414.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433–36.

Day, M. and Loffler, G. (2009). The role of orientation and position in shape perception. *Journal of Vision*, **9**(10), 14.1–17.

Dickinson, J. E., Bell, J., and Badcock, D. R. (2013). Near their thresholds for detection, shapes are discriminated by the angular separation of their corners. *PLoS ONE*, **8**(5), 1–9.

Dickinson, J. E., Cribb, S. J., Riddell, H., and Badcock, D. R. (2015). Tolerance for local and global differences in the integration of shape information. *Journal of Vision*, **15**(3), 21–21.

Dickinson, J. E., Haley, K., Bowden, V. K., and Badcock, D. R. (2018). Visual search reveals a critical component to shape. *Journal of Vision*, **18**(2), 2–2.

Field, D. J., Hayes, A., and Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research*, **33**(2), 173–93.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2017). Global processing of random-phase radial frequency patterns but not modulated lines. *Journal of Vision*, **17**(9), 18–18.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018). Convergent evidence for global processing of shape. *Journal of Vision*, **18**(7), 7–7.

Habak, C., Wilkinson, F., Zakher, B., and Wilson, H. R. (2004). Curvature population coding for complex shapes in human vision. *Vision Research*, **44**(24), 2815–23.

Habak, C., Wilkinson, F., and Wilson, H. R. (2006). Dynamics of shape interaction in human vision. *Vision Research*, **46**(26), 4305–20.

Habak, C., Wilkinson, F., and Wilson, H. R. (2009). Preservation of shape discrimination in aging. *Journal of Vision*, **9**(12), 18.1–8.

Hess, R. F., Wang, Y. Z., and Dakin, S. C. (1999). Are judgements of circularity local or global? *Vision Research*, **39**(26), 4354–60.

Hess, R. F., Hayes, A., and Field, D. J. (2003). Contour integration and cortical processing. *Journal of Physiology-Paris*, **97**(2–3), 105–119.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**(3), 65–70.

Jeffrey, B. G., Wang, Y. Z., and Birch, E. E. (2002). Circular contour frequency in shape discrimination. *Vision Research*, **42**(25), 2773–9.

Judd, C. M., Westfall, J., and Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, **103**(1), 54–69.

Kenward, M. G. and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, **53**(3), 983–997.

Kurki, I., Saarinen, J., and Hyvarinen, A. (2014). Investigating shape perception by classification images. *Journal of Vision*, **14**(12), 1–19.

Li, W., Thier, P., and Wehrhahn, C. (2000). Contextual influence on orientation discrimination of humans and responses of neurons in v1 of alert monkeys. *Journal of Neurophysiology*, **83**(2), 941–954.

Loffler, G., Wilson, H. R., and Wilkinson, F. (2003). Local and global contributions to shape discrimination. *Vision Research*, **43**(5), 519–530.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysocs: transforming numbers into movies. *Spatial Vision*, **10**, 437–42.

Poirier, F. J. and Wilson, H. R. (2006). A biologically plausible model of human radial frequency perception. *Vision Research*, **46**(15), 2443 – 2455.

Poirier, F. J. and Wilson, H. R. (2007). Object perception and masking: Contributions of sides and convexities. *Vision Research*, **47**(23), 3001–3011.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schmidtmann, G., Kennedy, G. J., Orbach, H. S., and Loffler, G. (2012). Non-linear global pooling in the discrimination of circular and non-circular shapes. *Vision Research*, **62**, 44–56.

Wang, Y.-Z. and Hess, R. F. (2005). Contributions of local orientation and position features to shape integration. *Vision Research*, **45**(11), 1375–83.

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, **18**, 292–97.

Wilder, J., Fruend, I., and Elder, J. H. (2018). Frequency tuning of shape perception revealed by classification image analysis. *Journal of Vision*, **18**(8), 9–9.

Wilkinson, F., Wilson, H. R., and Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, **38**(22), 3555–68.

Wilson, H. R. and Wilkinson, F. (2002). Symmetry perception: a novel approach for biological shapes. *Vision Research*, **42**(5), 589–97.

Wilson, H. R., Wilkinson, F., Lin, L. M., and Castillo, M. (2000). Perception of head orientation. *Vision Research*, **40**(5), 459–72.

Chapter 3

Phase-selective masking with radial frequency contours

3.1 Abstract

Sensitivity to changes in the shape of a closed-contour figure is affected by surrounding figures (Vision Research 44 (2004) 2815-2823). We examined how between-contour masking depends on radial frequency. Experiment 1 replicated previous studies that found that masking between adjacent radial frequency (RF) patterns was greatest when the two shapes were phase aligned, and that the magnitude of masking declined approximately linearly with increasing phase offsets. In addition, we found that the effect of phase offset on masking was very similar for RFs ranging from 3 to 8, a result that suggests that sensitivity to phase decreases with increasing radial frequency. Experiment 2 tested this idea and found that phase discrimination threshold for single cycles of curvature was approximately proportional to radial frequency. Experiment 3 showed that both curvature maxima and minima contribute to phase

dependent masking between RF contours. Together, Experiments 1-3 demonstrate that the strength of phase-dependent masking does not depend on RF, but is related to sensitivity for phase shifts in isolated contours, and is affected by both positive and negative curvature extrema. We discuss these results in relation to properties of curvature sensitive neurons.

3.2   Experiment 1

The perception of global shape of closed-contour figures relies on the extraction and integration of local contour features to represent larger curvature elements (Altmann *et al.*, 2003; Geisler *et al.*, 2001; Wang and Hess, 2005). Prior research has used Radial Frequency (RF) contours to study how the human visual system detects, extracts, and integrates curvature information to represent shape (Bell and Badcock, 2009; Bell *et al.*, 2009; Day and Loffler, 2009; Hess *et al.*, 1999; Jeffrey *et al.*, 2002; Kurki *et al.*, 2014; Loffler *et al.*, 2003; Wilkinson *et al.*, 1998). By modulating the radius of a circle in polar coordinates, RF contours can represent a variety of increasingly complex shapes while allowing researchers to control important stimulus parameters (Wilson and Wilkinson, 2002; Wilson *et al.*, 2000). Much research has focused on the detection of RF contours in isolation, but less attention has been devoted towards studying spatio-temporal interactions between contours. Because local features affect detection thresholds of other unrelated objects within naturalistic images (Alam *et al.*, 2014), studying the local and global interactions between shapes presumably is needed to better understand how the human visual system processes form in naturalistic contexts.

Previous work investigating spatial interactions between shapes defined by RF

contours has demonstrated that these interactions depend on the rotational phase alignment between a mask and target stimulus (Habak *et al.*, 2004, 2006, 2009). When the points of maximum curvature are aligned between RF contours (i.e., zero-phase difference), thresholds for detecting RF curvature are significantly elevated relative to a baseline condition that contains no masking stimulus, and the strength of masking diminishes as the relative phase difference increases between the mask and target RF contours (Habak *et al.*, 2004, 2006, 2009). These results are consistent with behavioural models that suggest shapes are discriminated using points of maximum orientation difference from circularity (i.e., curvature maxima) along a contour (Dickinson *et al.*, 2013, 2015, 2018), as offsets in spatial alignment of curvature maxima between two RF contours results in less masking. However, it is unclear how this relation between masking and phase alignment depends on the radial frequencies of the mask and target contours. Given that sensitivity to changes in curvature depends on radial frequency, and that low and high RF patters appear to be processed by different mechanisms (Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012), it is reasonable to suspect that phase-alignment masking would depend on the radial frequencies of the mask and target contours.

Previous studies suggest that summation of local curvature information along an RF contour depends on the frequency of curvature modulation (Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). For radial frequencies between 2 and 8 cycles/2π, detection thresholds improve with increasing number of modulation cycles at a rate that is greater than that predicted by probability summation (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). Based on these results, summation of curvature along the contour of these shapes is thought to rely upon the integration

of local information, possibly by a global shape mechanism (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). However, it is important to note that the rate of curvature summation differs across RF contours between 2 and 8 cycles/2π, which suggests that a global integration is more efficient for some radial frequencies compared to others (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). For high RF contours (i.e., > 8 cycles/2π), curvature summation is consistent with predictions of models that assume that the detection of curvature in extended contours is based on probability summation of local detectors (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012).

More recently, Baldwin *et al.* (2016) have questioned whether global pooling of local curvature is necessary to account for RF detection. By investigating RF detection under a Signal Detection Theory (SDT) framework rather than High Threshold Theory (HTT) as previous summation studies on RF detection have used, Baldwin *et al.* (2016) demonstrated that probability summation as a model describing the mechanism governing detection of RF contours cannot be rejected. Schmidtmann and Kingdom (2017) have extended this logic by developing a model that sufficiently accounts for curvature detection thresholds along both line and closed contour stimuli. A notable novelty in the model developed by Schmidtmann and Kingdom (2017) is the addition of a modulation transfer function for curvature frequency, which explains the flattening of detection thresholds to curvatures of intermediate and high frequencies, and aims to provide a unified theory for perception of curvature across contours.

Despite challenges to theories based on global summation of local curvature, this theory still remains plausible. Green *et al.* (2017, 2018a,b,c) conducted a series of

studies to examine how alterations to the original experimental design utilized by Baldwin *et al.* (2016) might influence the ability to distinguish between different theories of curvature summation across shapes. A key finding across these studies is that when phase of RF contours is randomized across trials, probability summation fails to predict performance across a wide range of experimental designs, even when assumptions of SDT are considered (Green *et al.*, 2017, 2018a,b,c). This finding suggests curvature frequency (i.e., the number of visible cycles of curvature) and *phase* may jointly determine the strength of summation of curvature around closed contours (Green *et al.*, 2017, 2018a,b,c).

Differences in the ability of shape mechanisms to integrate curvature along contours raises the possibility that the effect of phase alignment on masking may also vary as a function of radial frequency. Lateral interactions (i.e., masking) between RF contours are hypothesized to arise from the improper summation of curvature signals along target and mask contours at locations of peak curvature (Habak *et al.*, 2004; Poirier and Wilson, 2006). As the phase offset between a target and mask increases, the probability that curvature signals are improperly integrated across regions of peak curvature declines because the phase shift changes the polar angle at which these curvature maxima occur. As a result, masking declines with increasing phase offset between shapes (Habak *et al.*, 2004). Because summation of curvature signals depends on radial frequency and phase uncertainty, perhaps the phase-dependency of masking, thought to arise from improper summation of curvature signals at specific locations along a shape, also depends on radial frequency. This hypothesis predicts that shape interactions should operate over fixed distances, perhaps due to the fixed receptive field size of curvature sensitive neurons, which should cause the strength

66

of masking to vary with changes in physical distance between curvature extrema, rather than phase. Therefore, the effect of phase on masking should depend on radial frequency. However, if phase-dependent interactions between shapes instead depend upon the similarity between representations of a target and mask, as might occur, for example, if both contours were represented with a radial modulation function, then phase dependency may be constant across RF contours. Such a result would suggest that the summation that contributes to phase-dependent masking, and the type of summation that contributes to the detection of curvature along a shape, likely occur at different stages of visual processing.

The aim of the current study therefore was to investigate how the radial frequency of contour curvature affects the rate at which masking declines as a function of the phase alignment between two shapes.

3.2.1   Methods

Participants

One of the authors (MS) and eight naïve, experienced psychophysical observers participated in Experiment 1. Observer MS was tested in all conditions and the remaining observers were randomly assigned to one of four conditions, for a total of three observers per condition. The mean age of the observers was 20.8 years ($SD = 1.4$) and all observers had normal or corrected-to-normal visual acuity. Experimental protocols were approved by the McMaster University Research Ethics Board, and consent of the participant was collected prior to the start of the experiment.

Apparatus & Stimuli

Stimuli were presented on a Sony Model GDM-F520 monitor (1024 x 768) with a refresh rate of 100 Hz and a mean luminance of 62.9 cd/m^2. An APPLE G4 2.66 GHz Quad-Core Intel Xeon computer generated and displayed the shapes using Matlab 10.7.0 (Mathworks Inc., Natick, MA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). A chinrest was used to maintain a constant viewing distance of 131 cm. From this distance, a single pixel subtended 41.5 arc seconds.

Stimuli were comprised of RF contours generated by sinusoidally modulating the radius of a circle, and were defined by the equation (Wilkinson *et al.*, 1998):

$$r(\theta) = \bar{r}(1 + Asin(\omega\theta + \phi)) \tag{3.1}$$

where θ is the angle in radians, $\bar{r}$ is the radius of the contour, A is the amplitude of modulation expressed as a proportion of the radius of the circle, ω is the radial frequency in cycles per circumference (cy/2π), and ϕ is angular phase. The cross-sectional luminance profile of each contour was defined by a fourth derivative Gaussian (D4; see Wilkinson *et al.*, 1998). All RF contours had a luminance contrast of 99% and a peak spatial frequency of 8 cpd. All target RF contours had a mean radius of 1.14°. Each RF mask was positioned outside the target RF contour, and the mean target-mask distance was 0.35°. The target and mask always had the same radial frequency, and masking was measured with four radial frequencies (RF3, RF5, RF6, or RF8; see Figure 3.1). The modulation amplitude of each mask was set to 15 times the baseline detection threshold for that particular RF. The phase of the mask RF was always 0°, and thresholds were measured with seven target phases (0°, 30°, 60°, 90°, 120°, 150°, 180°). Examples of mask-target combinations are shown in Figure
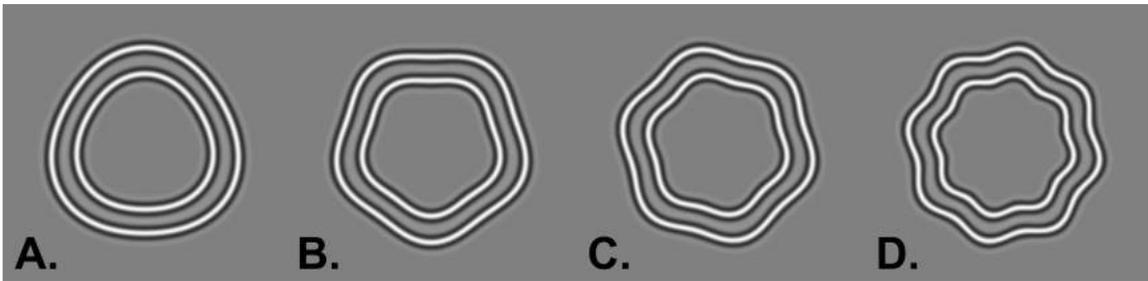
Figure 3.1: Example of stimuli used in Experiment 1. Stimuli were comprised of RF contours, here shown at a relative target-mask phase of 0°, with a contrast at 99% contrast and a deformation amplitude of 5%. The frequency of angular modulation per circumference varied across conditions: (a) RF3 (b) RF5 (c) RF6 (d) RF8.
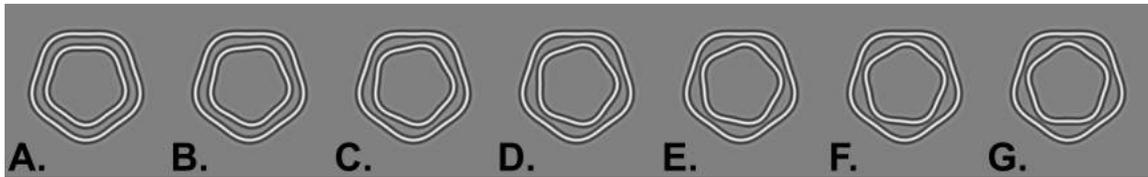
Figure 3.2: Example of seven target-mask phase combinations (A-G: 0, 30, 60, 90, 120, 150, & 180 deg) used throughout the study for an RF5 contour. The amplitude of both the target and mask are presented here at suprathreshold values for purposes of illustration only.

3.2. The four mask RFs and seven target-mask relative phases yielded a total of 28 masking conditions. For baseline, the phase of the target was varied across trials at three phase orientations 45°, 135°, 225°.

Procedure

Thresholds were measured with a two-alternative forced-choice paradigm and the method of constant stimuli. Each trial consisted of two stimulus intervals: one interval contained a circular contour (i.e., RF0), the other contained a target RF contour. Trials were initiated by observers by pressing the spacebar key on a computer keyboard. After a 300 ms interval a central fixation dot flickered for 500 ms and, after another 300 ms interval, two 150 ms stimulus intervals were presented with a 300 ms inter-stimulus-interval. The task was to select the stimulus interval that contained the target contour by pressing a key on a computer keyboard. No feedback was provided. The position of each contour on every trial was jittered within a 0.17° radius from the center of the screen; the mask and target were jittered in the same direction. Mean luminance remained constant throughout a testing session.

Prior to the start of the experiment, observers received three practice trials to familiarize themselves with the stimuli and task. Following practice, there was a 60 s light adaptation period, during which the observer fixated the center of the display. A single block of trials comprised either three or four target-mask phase combinations presented at seven different target modulation amplitudes, for a total of 504 or 672 trials per block, respectively. Each amplitude and phase combination was shown 24 times in a random order within a block. Each observer completed a block containing three different target-mask phase combinations (0°, 90°, 180°), and a block containing

four different target-mask phase combinations (30°, 60°, 120°, 150°) within a single experimental session. The order in which observers completed blocks was randomized within a session, and across observers. We adopted this block design to give observers a short break between testing blocks, and do not believe this decision in design influenced the outcome of our experiment. All observers completed two sessions, with each session being run on a separate day. In general, participants completed both sessions across two consecutive days, with each session lasting approximately 1.5 h.

Results

All analyses were performed using the statistical computing software R (R Core Team, 2017). Data for each target-mask phase combination were fit using maximum likelihood estimation with a psychometric function defined as:

$$\psi(x; \alpha, \beta, \gamma, \lambda) = \gamma + (1 - \gamma - \lambda)F_W(x; \alpha, \beta) \tag{3.2}$$

where x is RF modulation amplitude, γ and λ set the upper and lower asymptotes, and $F_W(x; \alpha, \beta)$ is a Weibull function (Weibull, 1951) defined as:

$$F_W = 1 - \exp(-(\frac{x}{\alpha})^{\beta}) \tag{3.3}$$

Threshold was defined as the RF modulation amplitude yielding 75% detection accuracy.

Thresholds were analyzed using a mixed linear model estimated with the lme4 package (Bates *et al.*, 2015). Degrees of freedom for the mixed model were approximated using the Kenward-Roger method (Kenward and Roger, 1997), which rescales the F ratios in addition to adjusting the degrees of freedom to better approximate

F-distributions for mixed linear models (Judd *et al.*, 2012). For brevity, we report only the F tests from the linear mixed-effects regression analyses (i.e., the analysis of variance of Type III sums of squares with Kenward-Roger approximation for degrees of freedom). Post-hoc comparisons between masking conditions were performed using paired, two-sided t tests with p values adjusted with the Holm-Bonferonni method and familywise $\alpha = 0.05$ (Holm, 1979).

Figure 3.3 displays RF detection thresholds averaged across all observers for each RF condition. For each radial frequency, threshold declined with increasing mask-target phase offset. In addition, close inspection of the figure indicates that detection thresholds decreased with increasing radial frequency. Finally, the effect of phase on threshold appeared to be very similar across RF conditions.

These set of observations were confirmed through the use of a mixed model, which was fitted with two fixed effects (RF condition and Phase) and one random effect (Observer). The ANOVA revealed significant main effects of RF condition ($F_{3,74.70} = 9.80$, $p < .0001$) and Phase ($F_{1,68.22} = 78.89$, $p < .0001$). The interaction between RF condition and Phase was not significant ($F_{3,68.22} = 2.45$, $p = .07$).

A more focused and potentially more powerful test of the RF $\times$ Phase interaction was conducted by calculating the linear and quadratic trends of threshold across phase offset (see Figure 3.4). The linear trend was consistently below zero and approximately equal for all RF conditions, whereas the quadratic trend was near zero in all RF conditions. These analyses suggest that thresholds decreased approximately linearly with phase offset at similar rates in all RF conditions.
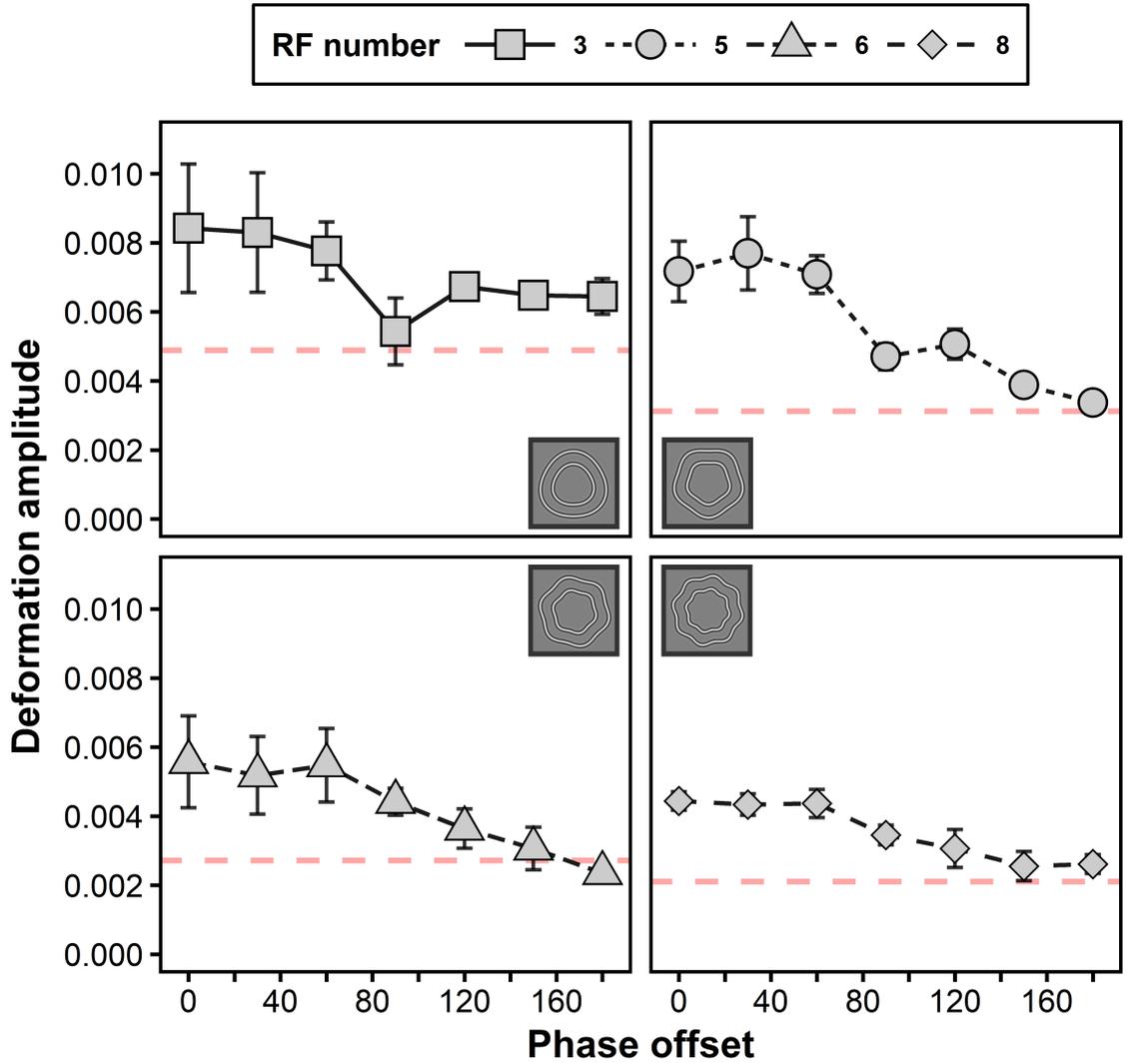
Figure 3.3: Detection thresholds as a function of the target-mask phase offset across the four RF conditions. In general, as the phase offset between the target and mask increases, there is a decrease in the amplitude of modulation required to make a correct detection judgement. Baseline measures for each RF contour are represented by the horizontal dashed red line. Error bars represent ±1 SEM.
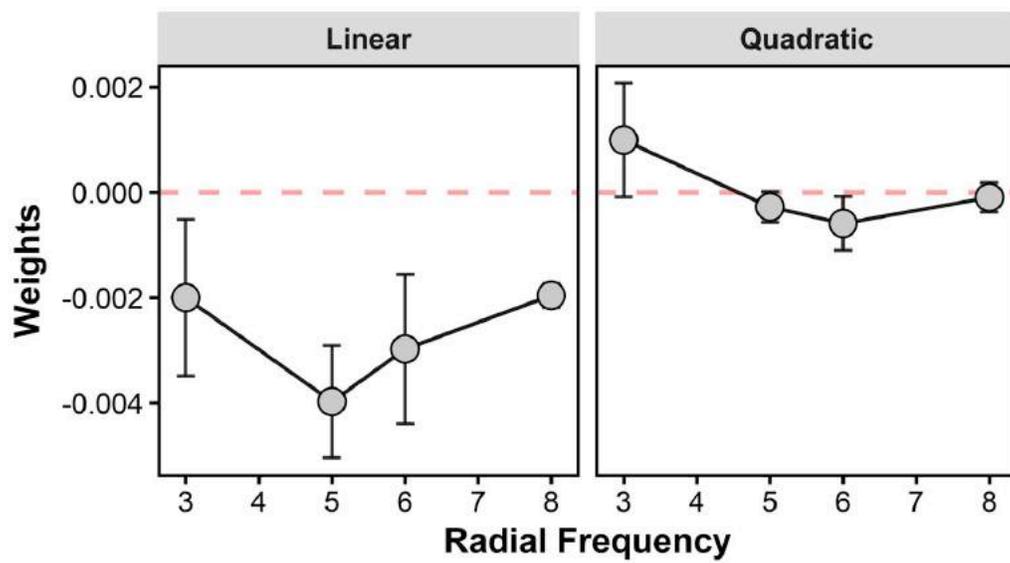
Figure 3.4: Values of the linear and quadratic trends of curvature detection thresholds across phase offsets plotted as a function of radial frequency. Error bars represent ± 1 SEM.

3.2.2 Discussion

Results from Experiment 1 demonstrate a decline in thresholds with increasing RF, consistent with findings from previous studies (Dickinson *et al.*, 2012; Schmidtmann and Kingdom, 2017; Wilkinson *et al.*, 1998). Furthermore, data from Experiment 1 show that phase alignment between target and mask contours affects the strength of masking. Lastly, the decline in threshold was approximately linear across phase offsets and the linear trend did not differ appreciably across RF conditions. These results are inconsistent with theories of masking between RF contours that hypothesize masking is caused by improper pooling of features across the receptive fields of curvature sensitive neurons (Habak *et al.*, 2004, 2006, 2009; Poirier and Wilson, 2007). Because these theories predict that masking should vary with the physical distance between curvature extrema, the effect of phase on masking should vary as a function of radial frequency, as an equal change in phase results in different changes in absolute polar angle across RF contours (see Figure 3.5). Furthermore, the constant effect of phase observed across RF conditions is surprising in light of previous work suggesting that curvature is summed differently along low- and high-frequency radial contours.

Perhaps current theories of masking between shapes cannot explain these findings because they do not take into consideration an observer's ability to detect relative phase shifts as a function of radial frequency. For example, the uniform effect of phase on RF masking is consistent with the idea that sensitivity to phase *per se* declines with increasing radial frequency. To our knowledge, the assumption that phase discriminability differs across RF contours has not been tested. Given its importance for explaining RF phase-dependent masking, Experiment 2 tested the assumption by measuring sensitivity for phase shifts of single-cycle RF contours as
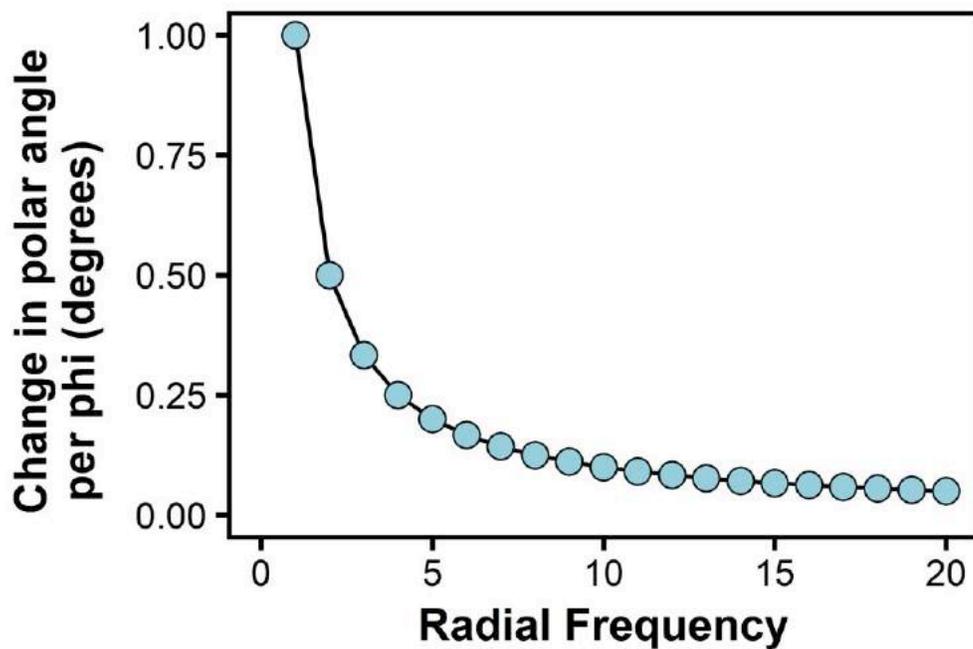
Figure 3.5: Figure displays the absolute change in polar angle that results from a single degree change in phase (ϕ) as a function of radial frequency. As radial frequency increases, changes in polar angle accompanying a shift in phase decline.

a function of radial frequency. Specifically, we predict that phase discrimination thresholds should be proportional to radial frequency.

3.3 Experiment 2: Does sensitivity to shifts in phase vary across radial frequencies?

Theories of masking between RF contours hypothesize that masking results from excess pooling of curvature features by curvature-sensitive neurons (Habak *et al.*, 2004, 2006, 2009; Poirier and Wilson, 2007), possibly in V4 (Pasupathy and Connor, 2001, 2002). These theories predict that sensitivity to shifts in phase may differ between RF contours as a consequence of the organization of receptive fields positioned radially from the origin of fixation. Because the range of polar angles over which a single cycle of modulation occurs is inversely related to radial frequency, neurons in V4 that respond to curvatures at different polar angles will produce responses that are more spatially dispersed for low RF contours, and more localized for high RF contours. Following a change in phase, which alters the polar angles of curvature features such as maxima and minima, RF contours that evoke dispersed responses may be more difficult to discriminate because the population response produced by V4 neurons will be similar across larger phase shifts compared to shapes that evoke highly localized responses. However, as mentioned in the Discussion section of Experiment 1, it is important to realize that the absolute change in polar angle produced by a constant phase shift decreases as radial frequency increases (see Figure 3.5). In other words, shifting curvature features (e.g., curvature maxima) by a specific polar angle requires a larger phase shift at high RFs than low RFs. According to this line

of reasoning, the similar results that we obtained across RFs (Figure 3.3) implies that sensitivity to phase shifts of RF patterns *decreases* with radial frequency. Therefore, for the effect of phase offset between a target and mask to remain constant across RF contours, as was found in Experiment 1, the theories of RF contour masking outlined above must assume that sensitivity to phase shifts of single cycles of curvature decreases as a function of radial frequency. Specifically, we predict that phase discrimination thresholds should be proportional to radial frequency. To our knowledge, this assumption has not been tested. Given its importance for explaining RF phase-dependent masking, Experiment 2 tested the assumption by measuring sensitivity for phase shifts of single-cycle RF contours as a function of radial frequency.

3.3.1   Methods

Participants

Five new naïve observers participated in Experiment 2. The mean age of the observers was 27.6 years ($SD = 1.9$) and all observers had normal or corrected-to-normal visual acuity. Experimental protocols were approved by the McMaster University Research Ethics Board, and consent of the participant was collected prior to the start of the experiment.

Apparatus & Stimuli

The apparatus was the same as the one used in Experiment 1. The stimuli were single cycle contours defined by the equation (adapted from Loffler *et al.*, 2003)

$$r(\theta) = \bar{r}(1 + B\frac{\theta - \theta_c}{\sigma}\mathrm{e}^{-(\theta-\theta_c)^2/\sigma^2}), \text{for} \quad \begin{aligned} \theta_c + \tfrac{2}{\omega}\pi \geq \theta > \theta_c \\ \theta_c - \tfrac{2}{\omega}\pi \leq \theta < \theta_c \end{aligned} \qquad (3.4)$$

$$= \bar{r} \quad \text{elsewhere}$$

Parameters of the first derivative Gaussian, B and σ, were constrained to those values that best matched the maximum slope, and maximum and minimum deviation in amplitude of a single cycle of modulation as sampled from a full RF pattern (as defined in Equation 3.1). To ensure that each cycle was clearly visible to observers, the maximum and minimum amplitude of each cycle was set to 15× the baseline detection threshold obtained with single cycle patterns. The radial frequency of a single cycle depends on where the upper and lower bounds are defined in Equation 3.4, which is determined by ω. Parameters $\bar{r}$ and θ for these single cycle RF contours were identical to those used in Experiment 1. Angular phase of patterns was determined by θ_c, which was defined by:

$$\theta_c = \frac{\phi}{\omega} \qquad (3.5)$$

where ϕ represents the phase of the pattern as defined in Equation 3.1. Therefore, ϕ changes the phase of a pattern by changing the value of θ_c, as the value of ω remains fixed for a given a pattern. In total, four RF contours (defined by ω) were tested: RF3, RF5, RF8 and RF11 (see Figure 3.6).
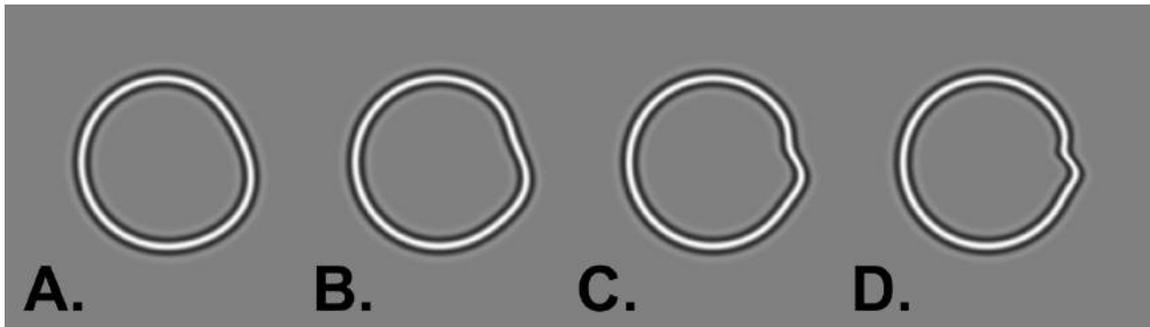
Figure 3.6: Example of single cycle RF contours used in Experiment 2 to test for sensitivity to shifts in phase. The amplitude modulation of each cycle was set at 15x the baseline detection threshold for that RF condition for each observer. In total, three RF conditions were tested: (a) RF3 (b) RF5 (c) RF8 (d) RF11. The amplitude modulation of each cycle was set to 0.07 in this figure to make undulations readily noticeable, while in the experiment amplitude modulations never reached such large values. All patterns are shown at θ_c set to zero.

Procedure

Phase discrimination thresholds were measured with a two-interval forced choice procedure. On each trial, two shapes were flashed onscreen across two intervals: each interval had a 150 ms duration, and the inter-stimulus interval was 500 ms. The first shape had an initial starting position (θ_c) set to a random value between 0 and 2π. The second shape had a phase of $\phi + \delta_\phi$, where ϕ was the same pedestal phase as the first, and δ_ϕ was a shift in phase. The observer's task was to determine if the second pattern was rotated clockwise or counterclockwise relative to the first pattern. The phase shift (δ_ϕ) was adjusted across trials with two independent, inter-leaved staircases. Both staircases followed a 2-down, 1-up rule, that terminated after 16 reversals, with final thresholds being defined by the average phase shift discriminated by observers of the last 8 reversals. The staircase step size started at 0.2 log units of phase shift from an initial starting value of 0.1066 radians (approximately 6.1 degrees), and was gradually reduced to a final step size of 0.02 log units following 2, 5, 8, and 11 reversals. On half the trials, the change in phase was positive, while on the remainder of trials it was negative. Observers completed 20 practice trials prior to the start of the experiment to ensure they understood the task. Each observer completed a minimum of 2 sessions per RF condition.

3.3.2   Results

The results from Experiment 2 are shown in Figure 3.7. In every observer, phase shift thresholds were lowest for RF3 contours and increased monotonically with increasing radial frequency. Thresholds were analyzed with mixed linear models following the same procedures used in Experiment 2, with one fixed effect (RF condition)
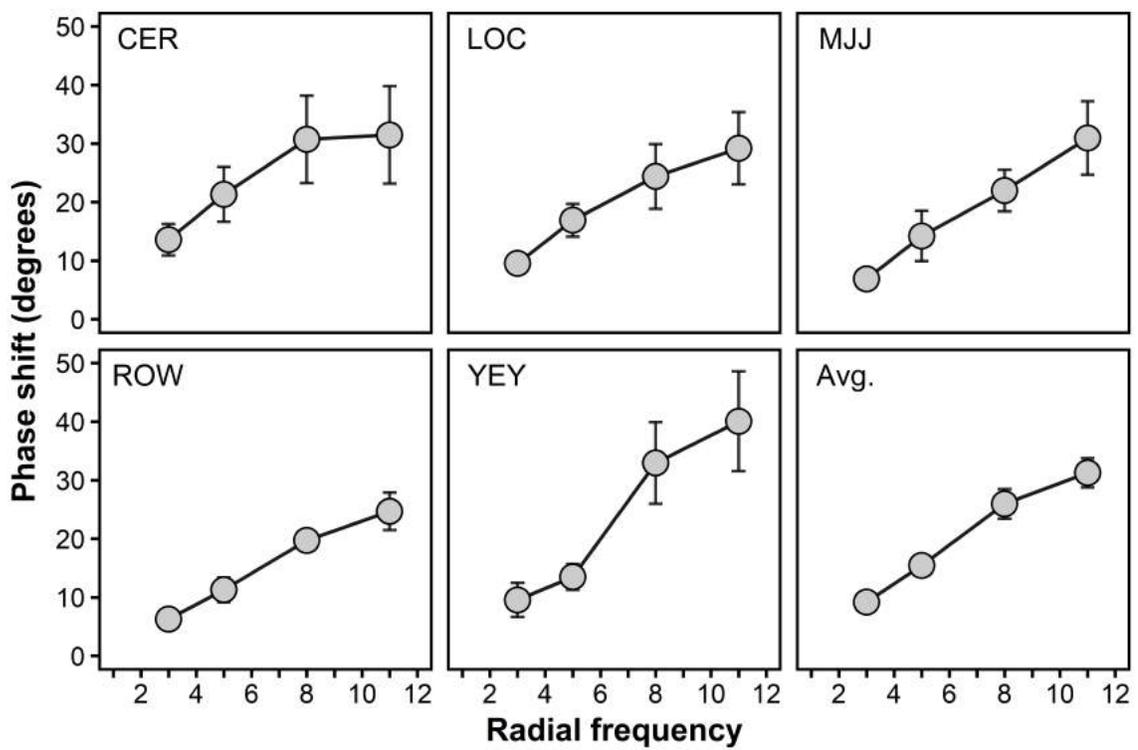
Figure 3.7: For each observer, plot of thresholds for phase shifts (degrees) for single cycle contours of four radial frequencies. Error bars represent ±1 SEM.

and two random effects (Observer and Session). The ANOVA revealed a significant effect of RF condition ($F_{3,7} = 15.22$, $p = .0018$). Post-hoc pairwise comparisons revealed significant differences between all RF conditions ($p_{Holm} < .01$) except between RF3 and RF5 conditions ($p_{Holm} = .05$) and between RF8 and RF11 conditions ($p_{Holm} = .15$).

To examine whether phase discrimination thresholds are proportional to radial frequency, we computed the best-fitting regression line for each subject, and then calculated the mean and median slope across linear fits. As can be seen in Table 3.1, the mean and median slope across observers was 2.82 and 2.42, respectively. Using the average values of the regression intercept and slope to calculate phase (ϕ) discrimination thresholds for each RF (ω), and Equation 3.5 to convert phase to polar angle (θ_c), we estimate that the just discriminable polar angle for our stimulus conditions was 3.07 deg.

3.3.3 Discussion

Experiment 2 found that sensitivity to shifts in phase for single cycles of curvature differ as a function of the radial frequency from which the cycles were sampled. It is unlikely that these results reflect an inability of observers to detect curvature deformations at different radial frequencies, as the amplitude of each RF contour was set to 15 times the curvature detection threshold in each condition.

Another possibility is that observers may have had greater trouble localizing the single RF cycle for high RF contours compared to low RF contours. Such differences in spatial uncertainty for RF location would make it difficult for observers to make comparisons of phase across stimulus intervals. Loffler *et al.* (2003) showed that,

Table 3.1: Intercept, slope, and R^2 values of best-fitting line for each observer's data, and for average data, from Experiment 2.

Observers	Intercept	Slope	R^2
CER	8.85	2.28	0.83
LOC	3.62	2.42	0.96
MJJ	-1.38	2.95	0.99
ROW	-0.34	2.34	0.98
YEY	-3.92	4.14	0.94
Mean	1.37	2.82	0.94
Median	-0.34	2.42	0.96
Averaged	1.37	2.82	0.97

relative to phase randomized conditions, fixing the phase of single-cycle RF patterns lowered curvature detection thresholds for intermediate RF contours (RF5) compared to low (RF3) or high (RF24) patterns. These results suggest that at least for the detection of single-cycles of curvature, spatial uncertainty has a greater effect on performance for intermediate than for low or high RF contours. This non-monotonic effect differs from the linear relationship that we found between phase discrimination thresholds and RF, which suggests that spatial uncertainty is unlikely to be the primary reason why phase discrimination thresholds decline with increasing radial frequency. Of course it is not clear that the differences in spatial uncertainty across radial frequencies observed by Loffler *et al.* also extend to the stimulus conditions used here, in which RF amplitude was well above detection threshold. Therefore, it remains possible that the results in Figure 3.7 may be due, at least in part, to the effects of spatial uncertainty.

A third possibility is that observers use absolute change in angular position of curvature maxima to detect changes in phase, much like observers use these points to discriminate between RF contours (Dickinson *et al.*, 2013, 2015, 2018). In the context of phase discrimination, an observer could locate a point of maximum curvature, and track the change in position of that point regardless of the curvature occupying the remainder of the contour. This explanation accurately accounts for our results, but does not explain the neural mechanism being used to encode shape. Therefore, we also propose that that this encoding scheme reflects the organization and size of curvature sensitive cells in V4 (Habak *et al.*, 2004, 2009). The overlap between receptive fields in V4 produces a population response, like that modeled by Pasupathy and Connor (2002), that is more discriminable for phase shifts of lower RF contours compared to

phase shifts for higher RF contours. Indeed, equivalent phase shifts produce larger angular rotations of shape for low RF patterns, as seen in Figure 3.5. Therefore, a decision making process monitoring the output of such a population code could make more acute judgments of phase for low versus high RF contours. Given such theories also predict that responses from curvature sensitive cells are integrated across curvature maxima (Dickinson *et al.*, 2013, 2015, 2018; Habak *et al.*, 2004; Poirier and Wilson, 2006), in Experiment 3 we investigate the importance of both maximum and minimum curvatures in contributing to effects of phase-dependent masking between shapes.

3.4 Experiment 3: What contribution do maxima and minima of curvature have on phase dependent masking?

Models of shape perception often use the angular separation between curvature maxima and/or minima on closed-contour shapes to generate scale invariant representations of form (Dickinson *et al.*, 2013, 2015, 2018; Pasupathy and Connor, 2001, 2002; Poirier and Wilson, 2006). Furthermore, evidence from psychophysical studies suggest that the extent of alignment between curvature maxima (and minima) influences the magnitude of masking observed between two shapes (Habak *et al.*, 2004, 2006, 2009; Poirier and Wilson, 2007). Experiment 3 therefore investigated the contribution of maxima and minima of curvature along a mask in modulating the effect of phase offset on masking strength that was observed in Experiment 1.

3.4.1   Methods

Participants

Three young adults participated in Experiment 3 ($M = 23.33$ years; $SD = 3.05$). Observers MS and CR participated in Experiments 1 and 2, while observer AS was new to the study. Observer MS was the author, and observers AS and CR were naïve to the purpose of the experiment. All observers had normal or corrected-to-normal visual acuity. Experimental protocols were approved by the McMaster University Research Ethics Board, and consent of the participant was collected prior to the start of the experiment.

Apparatus & Stimuli

The apparatus was the same as the one used in Experiments 1 and 2, with mask amplitude set to $15\times$ the baseline detection threshold for full cycle patterns. To isolate curvature extrema along masks, RF contours were positive or negative half-wave rectified yielding stimuli that contained only one kind of curvature feature (maxima or minima). Normal, unrectified RF contours were the same as those used in Experiment 1. The rectified contours contained positive and negative curvatures of unequal amplitude (i.e., stronger positive or negative component), whereas the full RF contour contained positive and negative curvatures of equal amplitude. It is important to note that half-wave rectified contours will contain curvatures of the opposite polarity at the junctions where a cycle rejoins the circular contour. For example, a convex half cycle will contain concavities on both sides of its deformation. This is a problem only if the operations that contribute to masking between shapes use unsigned

curvature signals, as Max-Min extrema are perfectly aligned at 90° phase offset between half-wave rectified masks and full wave targets. However, we see improvements in thresholds moving from 0° to 90° phase offset between target and masks across all RF masking conditions (see Results section for Experiment 3), at least partially suggesting that curvature signals are encoded using a signed representation.

Examples of stimuli from Experiment 3 are shown in Figure 3.8. A RF5 contour was used because Experiment 1 found that the influence of phase was strongest at that radial frequency (Figure 3.3).

Procedure

The procedure was identical to that used in Experiment 1, except that target-mask phase offsets were restricted to only three values: 0°, 90°, and 180°.

3.4.2   Results

Figure 3.9 displays detection thresholds for Experiment 3. When a target and mask were aligned in phase, thresholds were highest in the full contour condition, slightly lower in the condition in which masks contained only maximum (i.e., convex) curvature, and lowest in the condition where they contained only minimum (i.e., concave) curvatures. Performance was similar across conditions when target and masks were out of phase (i.e., phase offset = 180°).

Thresholds were analyzed with mixed linear models and the same corrective procedures used in the previous experiments. The model included two fixed effects (Curvature condition and Phase) and two random effects (Observer and Session). Quantitative analyses were consistent with observations described above: There were
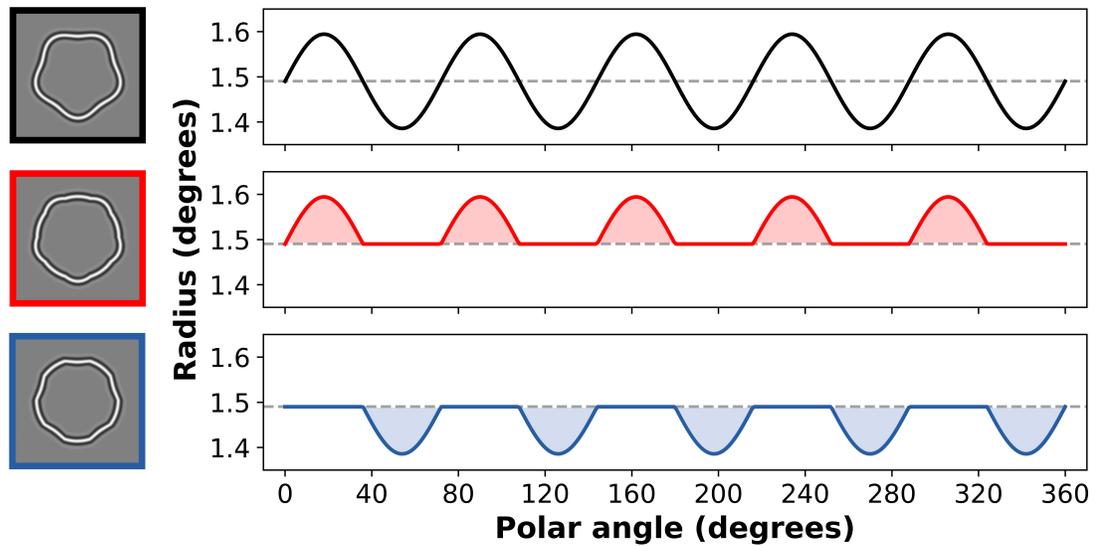
Figure 3.8: Examples of full and half-wave rectified RF5 stimuli used in Experiment 3 (left panels) along with plots demonstrating how the radius changes as a function of polar angle (right panels). *Top:* Full wave RF5 contour. *Middle:* Positive half-wave rectified RF5 contour containing only positive curvature segments. *Bottom:* Negative half-wave rectified RF5 contour containing only negative curvature segments.
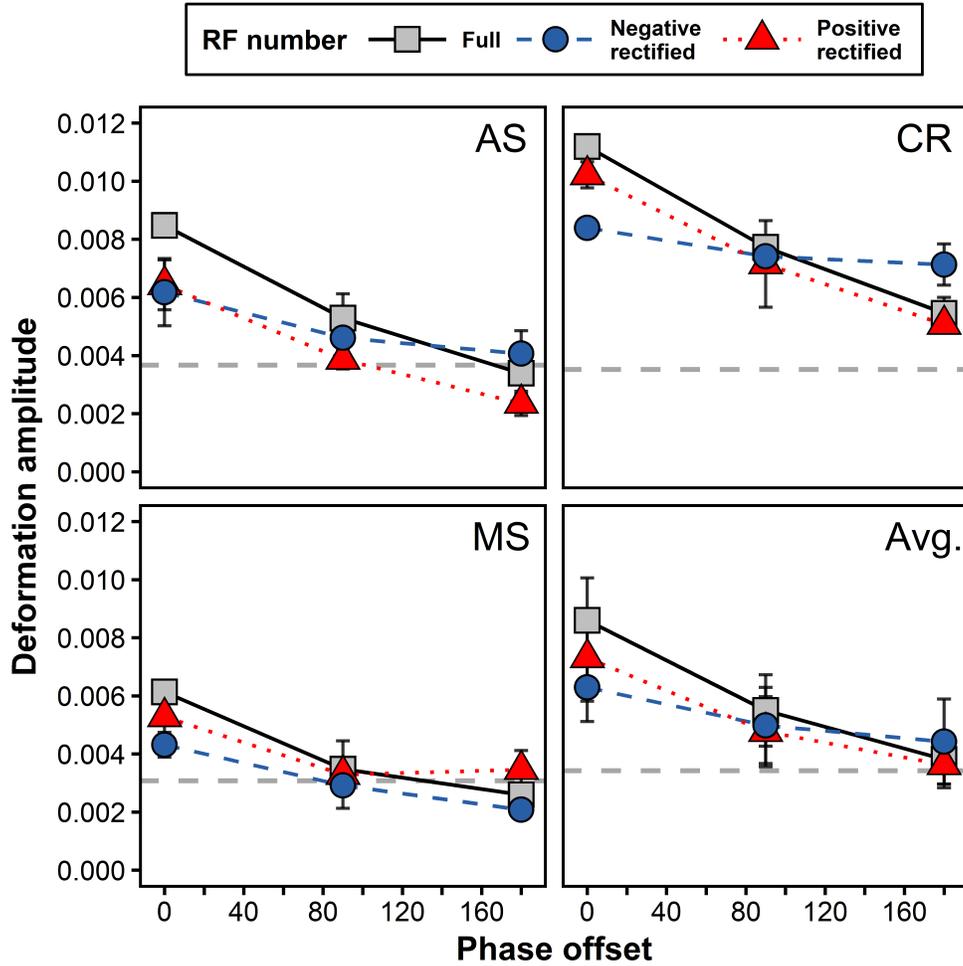
Figure 3.9: Curvature detection thresholds measured in Experiment 3. Thresholds for full and half-wave rectified RF contours are represented by separate line types, and baseline thresholds are indicated by the horizontal dashed lines. Error bars represent ±1 SEM.

significant main effects of Curvature condition ($F_{2,45.08} = 4.74$, $p = .013$) and Phase ($F_{2,45} = 79.29$, $p < .0001$), and a significant Curvature condition $\times$ Phase interaction ($F_{4,45} = 5.70$, $p < .0001$).

Post-hoc comparisons between masking conditions were performed using paired, two-sided t tests with p values adjusted with the Holm-Bonferonni method and familywise $\alpha = 0.05$ (Holm, 1979). Comparisons were performed between curvature conditions at 0° phase, as this phase offset resulted in the greatest elevation in thresholds. The comparisons revealed significant differences between the full contour condition and both minimum ($p_{Holm} < .0001$) and maximum ($p_{Holm} = .0014$) curvature conditions. Also, there was a significant difference between maximum and minimum curvature conditions ($p_{Holm} = .014$).

Discussion

Experiment 3 found that the effect of target-mask phase offset on detection thresholds varied across curvature conditions. The effect of the mask was greatest when its contour contained both maximum and minimum curvatures, and its effect was smaller when only maximum curvatures were present, and even smaller when only minimum curvatures were present. Furthermore, in all curvature conditions, detection thresholds were highest when the target and mask were phase aligned (0°) and declined monotonically with increasing phase offset. These results are consistent with previous work showing that masking between RF contours is driven primarily by the alignment of curvature maxima between neighbouring contours (Habak *et al.*, 2004; Poirier and Wilson, 2006, 2007). Inspection of Figure 3.9 shows that masking in the phase-aligned condition was greater for mask contours that contained curvature

maxima rather than curvature minima, but masking was greatest when the mask contour contained both curvature maxima and minima. These results are inconsistent with previous suggestions that masking between contours is *exclusively* a consequence of alignments of curvature maxima (Habak *et al.*, 2004; Poirier and Wilson, 2006). However these findings are consistent with a model of RF detection whereby observers detect curvature based on the difference between points of maximum and minimum extrema around a shape, as less masking is observed when this cue differs between the mask and target contours (i.e., less masking in half-wave rectified conditions Schmidtmann and Kingdom, 2017). Overall, these results demonstrate the albeit weaker contribution of curvature minima to the effect of a mask on curvature detection thresholds.

3.5   General Discussion

These experiments examined how phase alignment affects the magnitude of interaction between two shapes. Results from Experiments 1 and 3 replicate previous work demonstrating that masking is at its peak when two shapes are phase aligned, and declines monotonically with increasing phase offset between shapes. Experiment 1 extends these results by showing that masking was (approximately) a linear function of the phase offset that, surprisingly, did not differ across radial frequency. Experiment 2 found that phase discrimination was best for low RF contours, with thresholds increasingly monotonically with radial frequency. Lastly, Experiment 3 demonstrated that alignments of *both* curvature maxima and minima contribute to phase dependent masking between RF contours, although curvature maxima contribute more to the strength of masking than do curvature minima.

Together, these results suggest that curvature frequency does not affect the strength of phase-dependent masking between shapes, but does influence observers' ability to discriminate phase shifts of single-cycle shapes. In the following section, we interpret these results in relation to theories of masking based on the organization of curvature sensitive cells in visual cortex, as well as consider alternative ideas where phase dependent interactions arise upstream from mechanisms that integrate curvature across closed-contours.

3.5.1   Shape masking and local orientation filters

Models of shape perception attribute interactions between shapes to two main factors: the parallel alignment of contour segments (Hess *et al.*, 1999; Poirier and Wilson, 2007), and alignment of curvature maxima relative to the center of a shape (Habak *et al.*, 2004; Poirier and Wilson, 2006, 2007). Shape masking arising from the parallel alignment of contour segments is hypothesized to originate in primary visual cortex via weak local inhibition between orientation selective filters (Poirier and Wilson, 2006, 2007). Inhibition between local, orientation-selective neurons results in reduced neural responses that encode local contour orientation, thus providing weaker signals to higher visual areas that integrate information to represent shape (Poirier and Wilson, 2007). If signal strength about local orientation along a contour is diminished, then more signal (i.e., more curvature) will likely be required for an observer to accurately detect changes in curvature along a shape, because detecting changes in curvature requires integrating local orientation along a contour (Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012).

However, local inhibition between orientation-selective filters likely contributes

only weakly to masking between shapes (Habak *et al.*, 2004; Poirier and Wilson, 2007). If masking between contours was mediated by orientation selective mechanisms in V1, then circular (RF0) masks should should produce significant masking. This prediction arises from the fact that, at detection threshold, the orientation of contour curvature along target shapes is more parallel to a circular mask than a curvature-modulated mask. However, previous work examining the effect of different RF masks on detection thresholds shows that circular masks cause little to no masking (Habak *et al.*, 2004). Within the context of our study, this failure to see significant RF0 masking suggests that the decline in masking observed with increasing target-mask phase offset seen in Experiments 1 and 3 is only partially attributable to the divergent orientation structure accompanying phase misalignments between shapes. Instead, models of shape perception attribute interactions between shapes to largely be a consequence of the organization of curvature sensitive cells in V4 (Habak *et al.*, 2004; Poirier and Wilson, 2006, 2007).

3.5.2 Phase dependent masking and alignment of curvature maxima

Several investigators have proposed that intermediate visual areas represent object outlines using a sparse coding scheme based on the angular separation between curvature maxima relative to the object center (Cadieu *et al.*, 2007; Dickinson *et al.*, 2013, 2015, 2018; Schmidtmann *et al.*, 2015). Physiological data on the response properties of V4 neurons provides support for this polar-based coding scheme, as V4 neurons respond to curvature extrema relative to the center of a visual stimulus (Pasupathy and Connor, 2001, 2002). The radial organization of V4 neurons allows for a

scale invariant representation of shape, but it potentially could contribute to masking by concentric closed contour figures: if curvature extrema in the concentric contours occur at the same polar angle relative to the center of a shapes, then the response of neurons may saturate, resulting in an impaired ability to encode changes in curvature. Visual masking studies have shown that greater masking is produced by two masks than one (Habak *et al.*, 2004), which is consistent with the idea that the responses of V4 curvature sensitive neurons may increase as more curvature features fall within their receptive fields. Habak *et al.* suggested that the response of curvature sensitive cells follow a compressive non-linear function with increasing number of curvature maximums falling in its receptive field. This non-linear response function implies that, as the mean response increases, a greater change in contour curvature will be required to produce a just noticeable change in the neuron's response. If observers use the responses of these V4 neurons to discriminate contour curvature, then thresholds will increase when a target contour is surrounded by a curved contour mask (Habak *et al.*, 2004).

In the framework proposed by Habak *et al.* (2004), phase-dependent masking is a consequence of curvature maxima falling within the same receptive fields of curvature sensitive neurons (see Figure 3.10). The results from Experiments 1 and 3 are consistent with this idea: As the phase offset between two shapes increases, curvature maxima will increasingly fall within non-overlapping receptive fields, thereby reducing the response of curvature selective neurons. This reduction in response will, in turn, reduce the amount of curvature that is needed to detect a change in the target's curvature. In other words, curvature detection thresholds should be greatest when the target and mask are phase aligned and decline as the phase offset increases, which
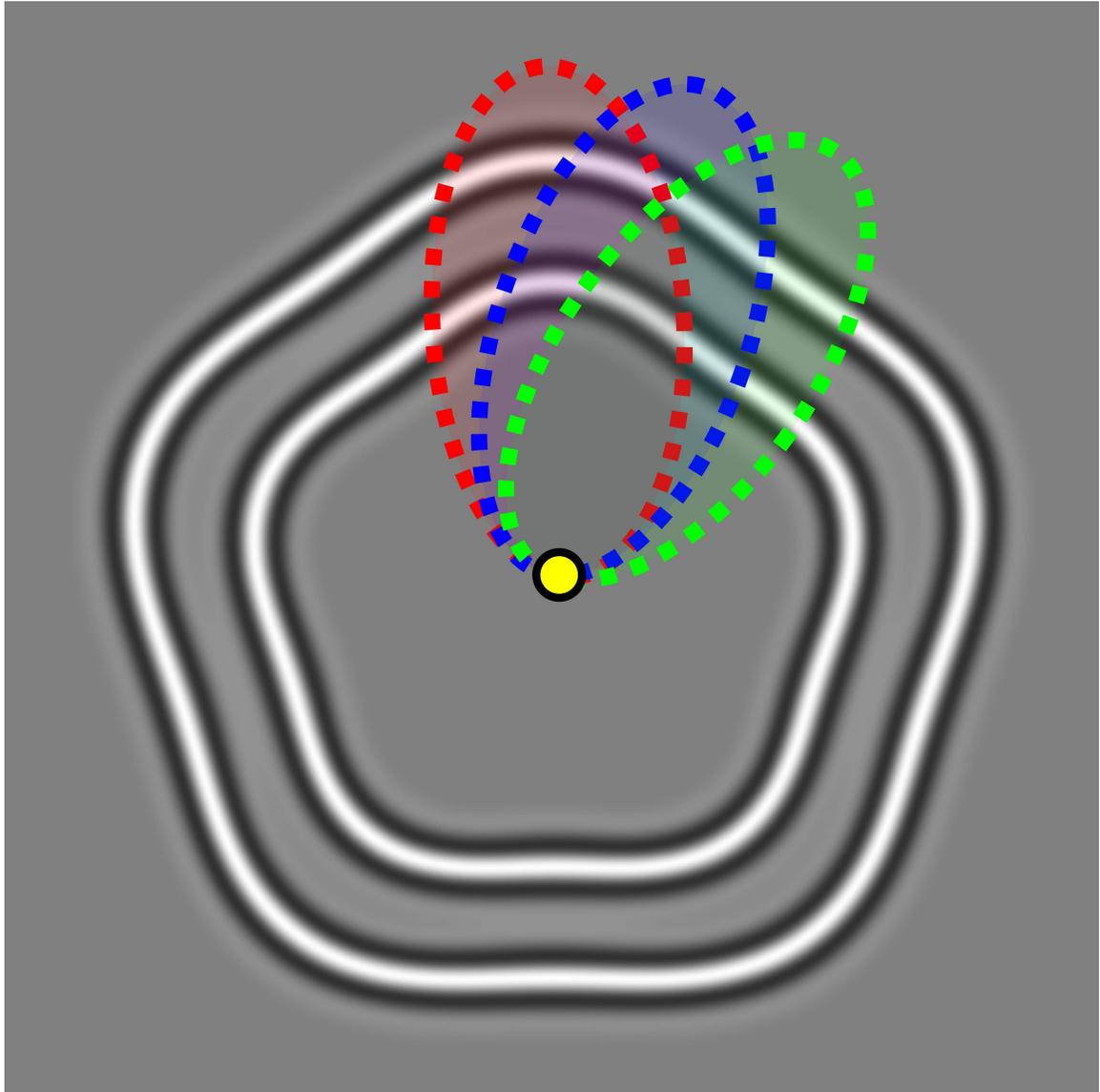
Figure 3.10: Example of receptive fields of curvature sensitive neurons, such as those in V4, organized relative to the center of a RF5 contour.

is consistent the results we obtained in Experiments 1 and 3 (see Figures 3.3 & 3.9).

Although excess pooling of contour curvature in V4 neurons explains why masking declines at greater target-mask phase offsets, it fails to explain why the effect of phase on thresholds is similar across RFs (Figure 3.3). The absolute change in polar angle produced by a constant phase shift decreases as radial frequency increases (see Figure 3.5). In other words, shifting curvature features (e.g., curvature maxima) by a specific polar angle requires a larger phase shift at high RFs than low RFs. If masking between shapes is entirely a consequence of excess spatial pooling of curvature features by V4 neurons, then smaller phase offsets should be required to significantly reduce masking in RF3 contours than RF11 contours unless, of course, the amount of excess/improper spatial pooling depends on *angle*, not *phase*. Contrary to this prediction, we found that the effect of phase offset on masking was similar across radial frequencies.

A possible explanation for our failure to find an effect of RF on phase-selective masking is that the effect of phase on masking between shapes occurs after the stage where local curvature signals are combined. Poirier and Wilson (2006) developed a biologically plausible model for RF detection whereby the output of intermediary curvature detectors are cross-correlated with patterns of activity for different shapes (e.g., sine-wave like) stored in memory. This process is akin to the visual system performing a Fourier-like decomposition of curvature frequency along a contour, and was computationally implemented as a Fourier transform by Poirier and Wilson. Perhaps the similarity between curvature frequencies of a target and mask are judged based upon the output of a process akin to computing the dot product between two vectors encoding shape. The output of such a process would be the same across RF conditions with shifts in phase, despite the absolute shift in polar angle with

changing phase differing as a function of radial frequency. These predictions also are consistent with previous studies arguing for the existence of shape channels, whereby different frequencies of curvature are encoded by channels sensitive to frequencies between approximately 3 - 10 cycles/2π (Bell and Badcock, 2009; Bell *et al.*, 2007, 2009; Habak *et al.*, 2004).

One argument against this idea is that even when shapes are phase aligned, if the distance between two identical shapes is great enough, no masking is observed (Habak *et al.*, 2004). This is one reason for arguing that masking between shapes is a consequence of curvature maxima falling within the same receptive fields of V4 neurons (Habak *et al.*, 2004). In a study using stimuli formed by strings of Gabors arranged along one of several RF contours presented at different radial distances, Schmidtmann *et al.* (2013) found that observers required fewer Gabors to detect modulated RF contours in noise when those Gabors were positioned along annuli of the same radius. Schmidtmann *et al.* argued that these results provide evidence for the existence of shape-specific analyzers that integrate orientation and position of local curvatures within individual annuli. Therefore, increasing the distance between shapes may result in a decline in masking because fewer curvature features (e.g., maxima and minima) are falling within the same V4 receptive field, and this may result in shapes being processed by separate shape analyzers.

Overall, our results from Experiment 1 are consistent with arguments based upon the existence of shape channels, suggesting that phase-dependent masking may arise *following* processes that integrate curvature around closed contour shapes.

3.5.3   Sensitivity to shifts in phase of curvature features

Experiment 2 found that sensitivity to phase shifts between contours varies with radial frequency: Observers were more sensitive to phase for low RF single cycle contours compared to higher RF single cycle contours (Figure 3.7). This increased sensitivity to phase is not the result of differences in curvature sensitivity across radial frequencies, as stimuli were presented at a constant multiple above detection threshold and, furthermore, the detection of contour curvature is phase (i.e., rotation) invariant (Bell and Badcock, 2009).

Another possible explanation is that the differences in sensitivity to phase result from the existence of multiple phase dependent channels whose tuning parameters differ across RF contours. Analogous to phase dependent masking, perhaps the ability to discriminate relative shifts in phase of shape relies upon a process that computes the similarity between stored representations of shape (i.e., shape channels), and the current shape being viewed. Because local curvature along low RF contours and internal representations of shape of similar frequency will be highly correlated despite changes in polar angle, such theories predict that larger angular rotations will be required for low RF contours to produce a just noticeable change in shape orientation as compared to high RF contours. In other words, such a theory predicts that phase sensitivity should remain the same across radial frequencies. However, such predictions are inconsistent with the results obtained in Experiment 2, that show observers' sensitivity to phase varying across radial frequency. Furthermore, it is unclear how this coding scheme would apply to shapes that contain large discontinuities in curvature along their contours (i.e., vertices), as such curvatures would contain power at extremely (infinitely) high frequencies.

Instead, differences in sensitivity to phase across RF contours may be a consequence of the population code used to encode shape. Behavioural data suggests that mechanisms used to code for shape identity might use a sparse representation based on angular positions of points of maximum curvature rather than dense representations based on the entire curvature around a shape (Dickinson *et al.*, 2013, 2015, 2018). Such a sparse encoding scheme is supported by physiological evidence that suggests outline of shapes are represented by the peak response from a population of neurons selectively tuned to both curvature and position (Carlson *et al.*, 2011; Pasupathy and Connor, 2001, 2002). In Experiment 2, curvature of single cycles of modulation differed as a function of the radial frequency from which they were sampled. Therefore, a population response of neurons sensitive to both curvature magnitude and position (i.e., polar angle), such as those in V4 (Carlson *et al.*, 2011; Pasupathy and Connor, 2001, 2002), will differ across RF contours even in the absence of shifts in phase of shape. Cycles of modulation sampled from low RF contours will have a large spatial spread, and thus elicit responses from neurons across many different polar angles. In contrast, curvature cycles sampled from high RF contours will elicit responses from neurons encoding a smaller range of polar angles. If the difference in population response evoked by shifts in phase is less informative for cycles of modulation sampled from certain RF contours compared to others, then the uniform effect of phase on masking may simply reflect this difference in sensitivity.

To test these predictions, and to better understand what information may be used by V4 to detect shifts of single cycle patterns from neural population codes, we developed a computational model that is similar to the one described by Pasupathy and

Connor (2002). The model, described in the following section, uses patterns of activation from neural population codes as a tool for formulating and evaluating linking propositions between behavioural data and the neurophysiology of shape perception. We briefly describe each stage of the model below, and emphasize that the simulation results that follow are not especially sensitive to minor changes in the values of hyperparameters chosen at each stage of the model.

Computational model of V4: Methods

The first stage of the model translates a shape from 2-dimensional Cartesian coordinates to polar coordinates by encoding the angular position and degree of radial modulation of the closed contour shape relative to object center. This stage is consistent with both behavioural (Habak *et al.*, 2004; Wilkinson *et al.*, 1998) and physiological (Pasupathy and Connor, 2001, 2002) evidence that suggests cortical areas along the ventral visual pathway may encode closed contour forms using such an encoding scheme.

The second stage of the model represents curvature (c), here defined as the rate of change in tangent angle with respect circular. We chose to use this definition of curvature, as this was the definition adopted in the model developed by (Poirier and Wilson, 2006), and was subsequently used by authors studying RF detection and discrimination (Dickinson *et al.*, 2013, 2015, 2018). For an alternative definition of curvature, see Schmidtmann and Kingdom (2017). Thus, we computed curvature as the second derivative of the contour's modulation amplitude, A, with respect to the shape radius (r) and polar angle, θ:

$$c_\theta = \frac{r^2 A_\theta}{r\theta} \tag{3.6}$$

Equation 3.6 cannot be solved analytically for single-cycle contours, and therefore we estimate curvature numerically using central difference differentiation.

Because abrupt changes in orientation (e.g., vertices) can result in discontinuities in otherwise smooth curvature functions, in Stage 3 the absolute curvature (c) at each point along the contour is replaced with a c' to mirror the measure used by Pasupathy and Connor (2001, 2002):

$$c'_\theta = \frac{2.0}{1 + \exp^{-\alpha c_\theta}} - 1.0 \tag{3.7}$$

Equation 3.7 is a sigmoidal function that ensures all curvature values along the contour are mapped onto values that range between -1 (strongly concave) to +1 (strongly convex). Parameter α, which controls the slope of the sigmoidal function, was set to 2 for our simulations because this value ensures that radial modulation amplitudes for the patterns in our experiments would elicit fairly strong responses from the neural population.

In the fourth stage of the model, values of c' are sampled at P equally-spaced points along the contour. In our simulations, P was set to an arbitrary value of 30, which was adequate for the contours used in our experiments. Changing P did not alter our key results except for low values that sparsely sample the c' function (i.e., $P < 8$).

Each curvature-sensitive unit (i.e., V4 neuron) responds to two stimulus parameters: position in polar angle (θ) and curvature (c'). In the fifth stage of the model,

the response of each unit is modelled as a 2-dimensional Gaussian function defined by Equation 3.8:

$$a_{map}(\theta, c') = \frac{1}{2\pi\sigma_\theta\sigma_{c'}}e^{-[((\theta-\mu_\theta)^2/2\sigma_\theta^2)+((c'-\mu_{c'})^2/2\sigma_{c'}^2)]} \qquad (3.8)$$

The shape of each 2-D Gaussian function for each unit was controlled by values of μ and σ, where μ determines the location of peak response (i.e., maximal activation) for the tuning function along the dimension under examination, and σ determines the spread of this function. We chose to model parameters that closely resembled that of Pasupathy and Connor (2002). Thus, the value of σ for angular position and curvature was 25° and 0.125, respectively.

In Stage 6, the overall response of a unit to a sampled stimulus with P points was set to equal the maximum response recorded for that unit in response to each point in P. This procedure is similar to that used by (Pasupathy and Connor, 2001), whereby the response r of a single unit is given by:

$$r(\theta, c') = \max_P[a_{map}(\theta, c')] \qquad (3.9)$$

The response of each unit will be strong if its preferred features are present within a shape, and will grade its response accordingly if those features deviate from this preferred combination of features.

In Stage 7, the response of each unit was scaled by a sigmoidal function that attenuated the response from units whose preferred curvature lies proximate to, or at zero. The function was defined as:

$$r_{weighted}(\theta, c') = \frac{1}{1 + e^{-a \cdot |c'|}} \cdot r(\theta, c') \qquad (3.10)$$

where a determines the slope of the function, and was set to a value of 2.5, which provides a modest attenuation of the response from units that prefer zero curvature. Applying such an attenuation to unit responses near zero curvature makes circularity a neural default, and is consistent with previous models of V4 (Carlson *et al.*, 2011; Habak *et al.*, 2004; Wilkinson *et al.*, 1998).

In Stage 8, we defined how many units N to model, which determines the density at which the feature space (defined by θ and c') is sampled. In total, 900 units were modelled that uniformly represented curvature across all angular positions, and approximated the number of units used by Pasupathy and Connor (2002).

Finally, responses at each combination of θ and c' are summed and normalized across all units:

$$S = p_{map}(\theta, c') = \frac{\sum_{k=1}^{N} r_{weighted}(\theta, c')}{\max[r_{weighted}(\theta, c')]} \qquad (3.11)$$

The final population response represented by matrix S is a surface that encodes changes in curvature along the contour of a shape at different polar angles.

Computational model of V4: Simulations

In total, eight simulations were run. Each single-cycle contour in Experiment 2 was run through the computational model to produce a response surface, where peaks along the surface represent maximal divergence in curvature (positive or negative) from circularity. Each contour was then rotated in polar angle by 3° and was

once again run through the model to produce a new population response. A 3° angle of rotation was chosen because this value is roughly equivalent to the minimum phase shift that could be detected for each single cycle contour tested in Experiment 2. Therefore, this angular rotation (i.e., differing phase shift across RF) would be especially diagnostic in understanding what factors limit the ability of observers in detecting shifts in phase of shape that contain different degrees of curvature along their contour. The difference in population response elicited from the presentation of each contour at a phase shift that was at threshold for human observers, as measured in Experiment 2, was then computed to examine what information may be available to the visual system to use to compute differences in shifts in phase across shapes.

Computational model of V4: Results

Results from the simulations are displayed in Figure 3.11. Examination of the population surfaces on the left on the figure demonstrates all three single-cycle contours elicit strong responses from curvature sensitive units, with the strength of activation of units identifying the presence of curvature extrema across polar angles. Restricting our examination along the curvature dimension also demonstrates that regions of peak activation for curvature extrema occur closer in feature space for low frequency single-cycle contours compared to contours sampled from higher frequency contours. Furthermore, the low-frequency contour elicits a broad pattern of activation across units, while intermediate and higher frequency shapes elicit more localized response patterns. Population surfaces generated by our model to shapes of different contour curvatures are consistent with the types of population responses to shape produced by Pasupathy and Connor (2001)'s model of V4 neurons.
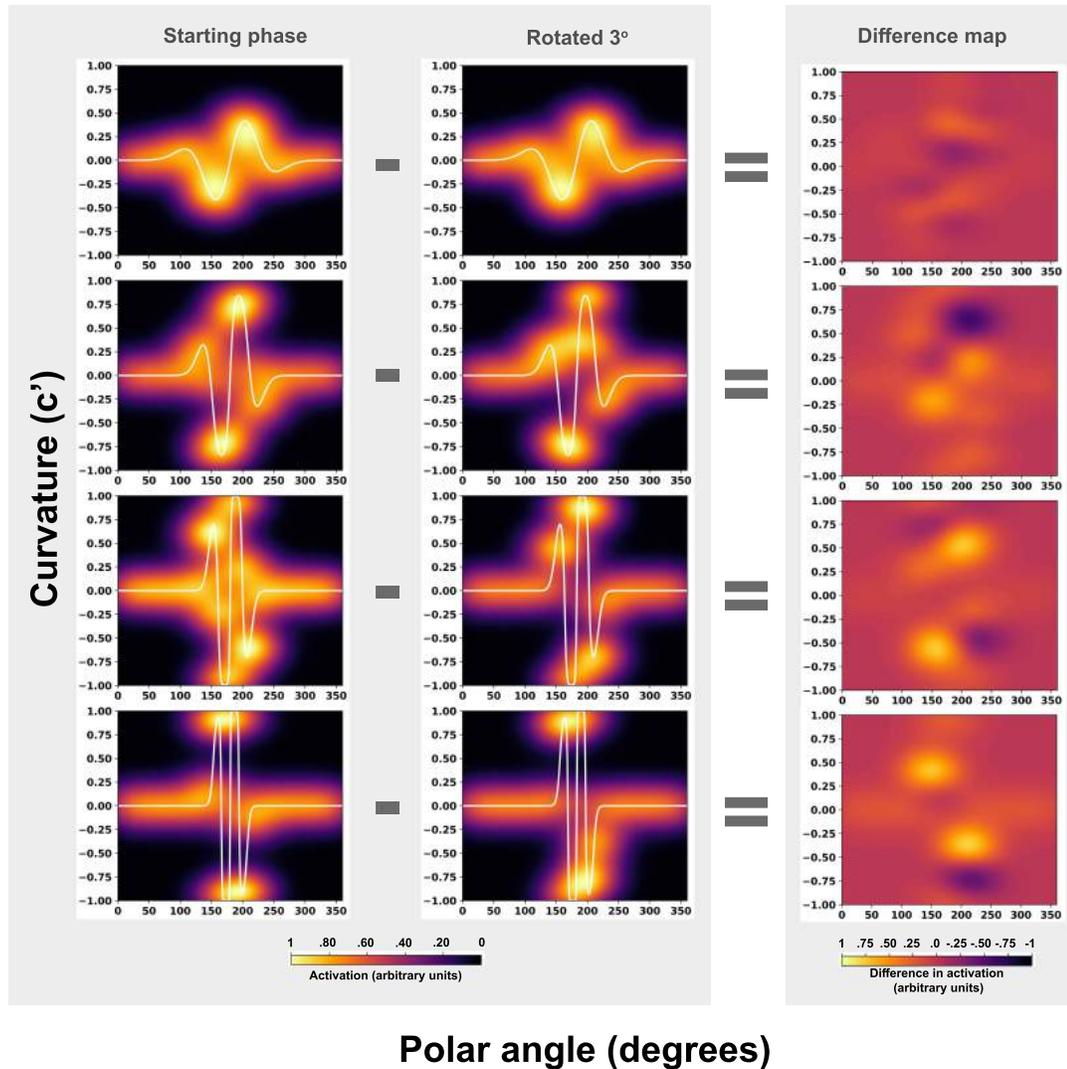
Figure 3.11: *Left & Middle columns:* Simulated population responses to contours used in Experiment 2. Unit activation ranged from 0 (black) to 1 (yellow). In each panel, the ordinate represents curvature, ranging from -1 (sharp concave) to 1 (sharply convex), and the abscissa represents angular position along the contour. The RF3, RF5, RF8, and RF11 contours used in Experiment 2 were used in all simulations: they are displayed in the bottom-left corner of each panel in the left column, and as white lines (c' mapping) in the left and middle columns. Responses were calculated for each RF at an initial angle of rotation and again after rotating the shape by 3°. The population surfaces generated did a good job at capturing curvature extrema, and more general curvature features across all polar angles. *Right column:* Difference maps produced by subtracting the response surfaces.

107

Despite a small angular rotation of 3° applied to each shape, the resultant population response was similar, but not identical, in units representing each version of the same shape at different polar angles. The difference map highlights what information may be utilized by the visual system in using such an encoding scheme to detect shifts in phase of shape. To quantify the similarity of population surfaces between phase shifted patterns, we computed the Mean-squared error (MSE) between activation maps generated by our model, as outlined in Table 3.2. Smaller MSEs suggest the surface maps are more similar, while larger MSEs suggest that surface maps are less similar. For the single-cycle RF3 contours, the population responses were very similar, resulting in a nearly flat difference map. These results suggest that the visual system may be utilizing sources of information besides curvature, or using a different scheme to encode curvature, to detect rotations of shapes that contain low-frequency curvatures. In contrast to low-frequency contours, difference maps generated for intermediate and high frequency contours contain noticeable differences in localized regions along the response surface. Therefore, information exists along the population surface that would allow the visual system to accurately discriminate rotations of these shapes. This result suggests that although such an encoding scheme can be used by the visual system in making such rotational judgments of shape that contain higher magnitudes of curvature, it is likely that other factors, such as noise introduced at various stages along visual processing, or spatial uncertainty may be responsible for limiting perception in these circumstances.

In Experiment 2, observers were more sensitive to shifts in phase of shape for low curvature contours compared to high curvature. Stated differently, sensitivity to angular rotation was roughly the same across RF contours. Given the difference

Table 3.2: Mean-squared error between activation surface maps generated by our V4 model when presented with a single-cycle RF contour, and the same contour shifted by 3° of angular rotation. Smaller MSEs suggest the surface maps are more similar, while larger MSEs suggest that surface maps are less similar.

Single-cycle RF number	Mean-squared Error (MSE)
RF3	0.018
RF5	0.029
RF8	0.034
RF11	0.036

map for the intermediate and high frequency sampled shapes both contain useful information in making a judgment as to whether the shape was phase shifted, why are thresholds for higher frequency shapes similar to that of the low frequency shape in Experiment 2? If performance at discriminating shifts in phase was correlated with the visual system detecting any difference in activity between population responses, then thresholds for detecting shifts in phase should be similar or *lower* for the high curvature shape compared to others tested. One possibility is that the population response generated in areas such as V4 is not the limiting factor in discriminating shifts in phase for shapes of different curvature. Perhaps, the visual system does not rely upon use of such representations to code for phase of shapes, or that curvature signals are altered at lower visual areas that further complicates the mapping of curvature signals onto a distributed population surface using such an encoding scheme.

An alternative explanation is that thresholds for the high frequency shape are similar to that for the low frequency shape because of the effects of spatial uncertainty on locating peaks of curvature for the single-cycle RF11 contour. Difference maps generated by our model for intermediate and high frequency sampled shapes contain diagnostic information to make judgments about the rotation of a shape. Therefore, what may be limiting performance in phase discrimination of shape at high radial frequencies may be the ability in coding the position of curvature extrema at specific polar angles. Such arguments are not immediately captured by the response properties of neurons in V4, as shown by our simulations. One way to make the difference maps more similar across RF contours in the model is by making the spread of the tuning curve that responds to location (i.e., polar angle), which corresponds to σ in Eq 3.8, increase as a linear function of radial frequency. This would have an overall effect

of introducing more noise in coding for location of curvatures depending on curvature frequency, and would help account for our results from Experiment 2. However, it is unclear whether spatial uncertainty arises as a result of changes in the properties of tuning functions that respond to curvatures at the level of V4, or whether this uncertainty may arise elsewhere in the visual system. Therefore, future work should aim to elucidate how spatial uncertainty may affect the population response of neurons in V4 at encoding properties of shape.

Computational model of V4: Comparison with other models

Many computational models of shape currently exist (Carlson *et al.*, 2011; Hatori *et al.*, 2016; Kempgens *et al.*, 2013; Pasupathy and Connor, 2001, 2002; Poirier and Wilson, 2006, 2010; Schmidtmann and Kingdom, 2017; Schmidtmann *et al.*, 2015). These models, as well as the one described here, build on the seminal work of Pasupathy and Connor (2001, 2002) suggesting that shape is encoded using a population surface that represents curvature at different polar angles.

Although the models differ significantly in their details, they all assume that neural response rates increase with the rate of curvature change, and the sparseness of the resulting representations depends on the gain of the response function. Hatori *et al.* (2016) demonstrated that by explicitly enforcing a penalty in their model to produce medium degrees of sparseness, curvature selectively and bias was produced in the population activity of units that resembled V4 neurons. Slugocki *et al.* (2018) have also shown that sparse representational vectors of shape can be used by machine learning agents to learn simple shape identities, which suggests that not all curvature information is pertinent to encoding identity. Attneave (1954) was one of the first

to suggest not all information encoded carries equal perceptual weighting; an idea inspired by information theory.

We expect that models that use a population code to represent curvatures (Carlson *et al.*, 2011; Kempgens *et al.*, 2013; Pasupathy and Connor, 2001, 2002; Poirier and Wilson, 2006), will be able to account for the effects of phase presented here. Specifically, the ability of a decision process analyzing these population responses to discriminate shifts in phase of shapes should be limited by the resolution at which the angular positions of curvature extrema are coded. However, our model implemented only a minor sparseness constraint in coding for curvature (i.e., Stage 7), and therefore the resultant population surface is able to capture more subtle differences in shifts in phase between two shapes. In contrast, models that enforce strong sparseness constraints (e.g., Carlson *et al.*, 2011; Hatori *et al.*, 2016) should require larger phase shifts to differentiate between complex shapes. Because extremely sparse coding schemes lose information about curvature occurring between curvature extrema, changes in the population response for these points cannot be used to judge shifts in phase between shapes. Thus, future work should examine how sparseness affects the ability of computational models to discriminate phase-shifted shapes of varying contour complexity. Furthermore, it would be interesting to investigate how other models of shape that do not rely upon position coding of curvature, such as Schmidtmann and Kingdom (2017), might predict human phase discrimination.

3.5.4 Importance of curvature extrema in masking between shapes

Masking studies using RF contours attribute the interaction between shapes to be primarily the consequence of the alignment of curvature maxima (Habak *et al.*, 2004, 2006, 2009; Poirier and Wilson, 2007), with little mention with regards to the contribution of curvature minima. Such theories of shape masking are also consistent with models of RF detection that claim only points of maximum curvature are used to code for shape identity (Poirier and Wilson, 2006). Contrary to previous work, results from Experiment 3 demonstrate that the alignment of either curvature maxima or minima result in masking between two identical, phase aligned shapes, although the effect of masking was greatest when the mask that contained both curvature extrema. One argument against this idea is that half-wave rectified contours still contain curvatures of the opposite polarity where the cycle rejoins the circular portion of the shape, albeit occurring at smaller amplitudes than in full waveform counterparts, and therefore masking can still occur between these features and full cycles appearing in the target. However, such arguments also predict that peak masking for the negative half-wave rectified condition should occur when the phase offset between target and masks is at 90°, as positive full-cycles of the target would be in perfect alignment with the positive curvatures occurring at the transitional segments in masks. Contrary to this prediction, less masking was found for 90° phase offsets relative to the condition in which both target and negative half-wave rectified contours are in perfect alignment (i.e., 0°).

Consistent with the results of Experiment 3, physiological evidence suggests the use of population codes within V4 to encode shapes that operate along a continuum of

features; namely curvature and angular position (Pasupathy and Connor, 2001, 2002). Therefore, although phase dependent masking is hypothesized to primarily result from the excess summation of curvature signals within V4 receptive fields, this explanation need not be exclusive to peaks of curvature. Instead, data from Experiment 3 argue that alignment of curvature minima may also produce effects associated with excess summation of curvature signals within receptive fields tuned to curvature. Data from Experiment 3 are also consistent with psychophysical studies that suggest both curvature types contribute equally to the perception of global shape (Bell *et al.*, 2010; Kempgens *et al.*, 2013; Schmidtmann and Kingdom, 2017; Schmidtmann *et al.*, 2015).

Data from Experiment 3 also are consistent with the model proposed by Schmidtmann and Kingdom (2017) whereby RF contours are detected based on the difference between maximum and minimum curvatures. If observers use curvature differences to detect changes in modulation amplitude, perhaps this metric might also be used by the human visual system to judge the similarity between shapes. Following this line of reasoning, greater masking should be observed between shapes that share a similar difference in curvature between Max-Min points. Therefore, less masking should be observed between a full-wave target and half-wave rectified masks in comparison to masks that contain a full waveform. Our data from Experiment 3 are consistent with such a model of curvature detection: less masking was observed in conditions where the Max-Min difference between targets and masks differed as a result of applying a half-wave rectification to masking contours. Future work should aim to extend the mathematical reasoning of this model to explicitly predict masking effects between shapes, and how phase-dependent masking might emerge using a difference score between Max-Min curvatures.

3.5.5   Conclusion

To summarize, we demonstrated that the effect of phase on masking between shapes does not differ across radial frequencies. However, sensitivity at discriminating shifts in phase does differ as a function of radial frequency, and helps to explain the uniform effect of phase on masking across RF contours. Surprisingly, differences in sensitivity to phase are not captured by models of V4 neurons that produce population surfaces in response to the magnitude and position of curvature, and therefore alternative explanations are needed. Finally, using half-wave rectified masks, we showed that both curvature maxima and minima contribute to phase dependent masking between shapes based on angular alignment of these features relative to the origin of a shape.

3.6 References

Alam, M., Vilankar, K. P., Field, D. J., and Chandler, D. M. (2014). Local masking in natural images: A database and analysis. *Journal of Vision*, **14**(22), 1–38.

Altmann, C. F., Bülthoff, H. H., and Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology*, **13**(4), 342–349.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, **61**(3), 183–193.

Baldwin, A. S., Schmidtmann, G., Kingdom, F. A. A., and Hess, R. F. (2016). Rejecting probability summation for radial frequency patterns, not so Quick! *Vision Research*, **122**, 124–134.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48.

Bell, J. and Badcock, D. R. (2009). Narrow-band radial frequency shape channels revealed by sub-threshold summation. *Vision Research*, **49**(8), 843–50.

Bell, J., Badcock, D. R., Wilson, H., and Wilkinson, F. (2007). Detection of shape in radial frequency contours: independence of local and global form information. *Vision Research*, **47**(11), 1518–22.

Bell, J., Wilkinson, F., Wilson, H. R., Loffler, G., and Badcock, D. R. (2009). Radial frequency adaptation reveals interacting contour shape channels. *Vision Research*, **49**(18), 2306–17.

Bell, J., Hancock, S., Kingdom, F. A. A., and Peirce, J. W. (2010). Global shape processing: Which parts form the whole? *Journal of Vision*, **10**(6), 16–16.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433–36.

Cadieu, C., Kouh, M., Pasupathy, A., Connor, C. E., Riesenhuber, M., and Poggio, T. (2007). A Model of V4 Shape Selectivity and Invariance. *Journal of Neurophysiology*, **98**(3), 1733–1750.

Carlson, E. T., Rasquinha, R. J., Zhang, K., and Connor, C. E. (2011). A Sparse Object Coding Scheme in Area V4. *Current Biology*, **21**(4), 288–293.

Day, M. and Loffler, G. (2009). The role of orientation and position in shape perception. *Journal of Vision*, **9**(10), 14.1–17.

Dickinson, J. E., McGinty, J., Webster, K. E., and Badcock, D. R. (2012). Further evidence that local cues to shape in RF patterns are integrated globally. *Journal of Vision*, **12**(12), 16–16.

Dickinson, J. E., Bell, J., and Badcock, D. R. (2013). Near their thresholds for detection, shapes are discriminated by the angular separation of their corners. *PLoS ONE*, **8**(5), 1–9.

Dickinson, J. E., Cribb, S. J., Riddell, H., and Badcock, D. R. (2015). Tolerance for local and global differences in the integration of shape information. *Journal of Vision*, **15**(3), 21–21.

Dickinson, J. E., Haley, K., Bowden, V. K., and Badcock, D. R. (2018). Visual search reveals a critical component to shape. *Journal of Vision*, **18**(2), 2–2.

Geisler, W. S., Perry, J. S., Super, B. J., and Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, **41**(6), 711–724.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2017). Global processing of random-phase radial frequency patterns but not modulated lines. *Journal of Vision*, **17**(9), 18–18.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018a). Convergent evidence for global processing of shape. *Journal of Vision*, **18**(7), 7–7.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018b). The effect of spatiotemporal displacement on the integration of shape information. *Journal of Vision*, **18**(5), 4–4.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018c). Integration of shape information occurs around closed contours but not across them. *Journal of Vision*, **18**(5), 6–6.

Habak, C., Wilkinson, F., Zakher, B., and Wilson, H. R. (2004). Curvature population coding for complex shapes in human vision. *Vision Research*, **44**(24), 2815–23.

Habak, C., Wilkinson, F., and Wilson, H. R. (2006). Dynamics of shape interaction in human vision. *Vision Research*, **46**(26), 4305–20.

Habak, C., Wilkinson, F., and Wilson, H. R. (2009). Preservation of shape discrimination in aging. *Journal of Vision*, **9**(12), 18.1–8.

Hatori, Y., Mashita, T., and Sakai, K. (2016). Sparse coding generates curvature selectivity in V4 neurons. *JOSA A*, **33**(4), 527–537.

Hess, R. F., Wang, Y. Z., and Dakin, S. C. (1999). Are judgements of circularity local or global? *Vision Research*, **39**(26), 4354–60.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**(3), 65–70.

Jeffrey, B. G., Wang, Y. Z., and Birch, E. E. (2002). Circular contour frequency in shape discrimination. *Vision Research*, **42**(25), 2773–9.

Judd, C. M., Westfall, J., and Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, **103**(1), 54–69.

Kempgens, C., Loffler, G., and Orbach, H. S. (2013). Set-size effects for sampled shapes: experiments and model. *Frontiers in Computational Neuroscience*, **7**.

Kenward, M. G. and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, **53**(3), 983–997.

Kurki, I., Saarinen, J., and Hyvarinen, A. (2014). Investigating shape perception by classification images. *Journal of Vision*, **14**(12), 1–19.

Loffler, G., Wilson, H. R., and Wilkinson, F. (2003). Local and global contributions to shape discrimination. *Vision Research*, **43**(5), 519–30.

Pasupathy, A. and Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, **86**(5), 2505–2519.

Pasupathy, A. and Connor, C. E. (2002). Population coding of shape in area V4. *Nature neuroscience*, **5**(12), 1332—1338.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysocs: transforming numbers into movies. *Spatial Vision*, **10**, 437–42.

Poirier, F. J. and Wilson, H. R. (2006). A biologically plausible model of human radial frequency perception. *Vision Research*, **46**(15), 2443 – 2455.

Poirier, F. J. and Wilson, H. R. (2007). Object perception and masking: Contributions of sides and convexities. *Vision Research*, **47**(23), 3001 – 3011.

Poirier, F. J. A. M. and Wilson, H. R. (2010). A biologically plausible model of human shape symmetry perception. *Journal of Vision*, **10**(1), 9–9.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schmidtmann, G. and Kingdom, F. A. A. (2017). Nothing more than a pair of curvatures: A common mechanism for the detection of both radial and non-radial frequency patterns. *Vision Research*, **134**, 18–25.

Schmidtmann, G., Kennedy, G. J., Orbach, H. S., and Loffler, G. (2012). Non-linear global pooling in the discrimination of circular and non-circular shapes. *Vision Research*, **62**, 44–56.

Schmidtmann, G., Gordon, G. E., Bennett, D. M., and Loffler, G. (2013). Detecting shapes in noise: tuning characteristics of global shape mechanisms. *Frontiers in Computational Neuroscience*, **7**.

Schmidtmann, G., Jennings, B. J., and Kingdom, F. A. A. (2015). Shape recognition: convexities, concavities and things in between. *Scientific Reports*, **5**, 17142.

Slugocki, M., Sekuler, A. B., and Bennett, P. J. (2018). Effect of signal alteration on learning shape identity using sparse representations. *Conference on Cognitive Computational Neuroscience.*

Wang, Y.-Z. and Hess, R. F. (2005). Contributions of local orientation and position features to shape integration. *Vision Research*, **45**(11), 1375–83.

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of applied mechanics*, **18**, 292–97.

Wilkinson, F., Wilson, H. R., and Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, **38**(22), 3555–68.

Wilson, H. R. and Wilkinson, F. (2002). Symmetry perception: a novel approach for biological shapes. *Vision Research*, **42**(5), 589–97.

Wilson, H. R., Wilkinson, F., Lin, L. M., and Castillo, M. (2000). Perception of head orientation. *Vision Research*, **40**(5), 459–72.

Chapter 4

Isotropic Sensitivity to Curvature Deformations Along Closed Contours

4.1   Abstract

Human observers are exquisitely sensitive to curvature deformations along a circular closed contour (Hess *et al.*, 1999; Loffler *et al.*, 2003; Wilkinson *et al.*, 1998). Such remarkable sensitivity is attributed to the encoding scheme used by V4 neurons to represent curvature at different polar angles across the visual field (Pasupathy and Connor, 2001, 2002). Computational models of the response properties of V4 neurons often assume that sensitivity to curvature remains constant across all polar angles (Pasupathy and Connor, 2001, 2002), and therefore predict that curvature detection thresholds ought to be invariant across polar angles. We tested this prediction with a novel stimulus class we call Difference of Gaussian contours which allowed us

to independently manipulate the amplitude, angular frequency, and polar angle of curvature of a closed-contour shape. We found that that observers are uniformly sensitive to positive curvature deformations at all 24 polar angles tested, and this result holds despite changes in angular frequency and spatial uncertainty.

4.2   Introduction

The retinal image produced by a visual scene is processed by a hierarchy of visual processing stages that organize information into more structured elements of perception, such as shapes and objects (Felleman and Van, 1991; Riesenhuber and Poggio, 1999). For example, neurons in primary visual cortex (V1) encode local contour orientation (Hubel and Wiesel, 1959, 1968; Hubel *et al.*, 1978), whereas neurons in V2 combine local orientation information to represent angles and arcs (Anzai *et al.*, 2007; Hegdé and Essen, 2000; Ito and Komatsu, 2004). Neurons in areas V3 and V4 respond to increasingly complex stimulus relations, such as the curvature along closed contours (Gallant *et al.*, 1996; Pasupathy and Connor, 1999, 2001, 2002), and IT neurons are selective for stimuli in particular object-level categories, such as faces (Tanaka, 2003).

For object form to be well represented based upon the input from lower visual areas, regions diagnostic of object identity (e.g., curvature extrema) need to be adequately sampled despite an increase in convergence of visual information moving up the visual processing hierarchy. However, given the uncertainty of where maximally informative regions might appear along the contour of an object, representations of curvature should be well preserved across all polar angles to ensure that these regions are adequately sampled. Previous neurophysiological studies examining the tuning

properties of neurons in V4 of macaques have demonstrated that the response of these cells is often well described by models that *assume* curvature is sampled uniformly across polar angles (Pasupathy and Connor, 2001, 2002). A prediction that emerges from such an assumption is that observers may be uniformly sensitive to curvature deformations at all polar angles, an assumption that remains untested. This prediction is surprising when one considers that many aspects of vision are relatively poorer for oblique contours compared to horizontal and vertical contours: oblique effects have been found for grating acuity (Berkley *et al.*, 1975; Campbell *et al.*, 1966; Teller *et al.*, 1974), orientation discrimination (Appelle, 1972; Heeley *et al.*, 1997; Mansfield, 1974; Westheimer, 2003), motion perception (Dakin *et al.*, 2005; Gros *et al.*, 1998; Westheimer, 2003), and many other visual tasks (Appelle, 1972). The oblique effect is thought to reflect differences in the both the number and response properties of visual neurons that encode contours at oblique and cardinal orientations (Bonds, 1982; Furmanski and Engel, 2000; Li *et al.*, 2003; Wang *et al.*, 2003; Xu *et al.*, 2006). These findings suggest that it is at least plausible that observers are not uniformly sensitive to curvature at all polar angles, but rather are more sensitive to curvature at horizontal and vertical orientations compared to oblique orientations.

Previous psychophysical studies investigating visual sensitivity at detecting curvatures along the contour of shapes have typically randomized the polar angle at which deformities appear, or kept the spatial location of deformation along contours fixed (Baldwin *et al.*, 2016; Bell and Badcock, 2009; Bell *et al.*, 2009; Hess *et al.*, 1999; Jeffrey *et al.*, 2002; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998). Consequently, it is difficult to judge whether visual sensitivity to curvature

varies as function of polar angle, or what effect spatial uncertainty might on sensitivity as a function of polar angle. Schmidtmann *et al.* (2015) did observe observe a lower visual field preference for human observers when discriminating between two shapes based on differences in curvature, suggesting that human visual system might preferentially weight curvature information along objects depending on an objects' location in the visual field. However, Schmidtmann *et al.* (2015) shifted the spatial location of the entire shape within an observer's visual field, rather than systemically varying only the angular subtense of the deformity while keeping the origin of the shape fixed. In a series of consecutive studies aimed at investigating the effect of spatial uncertainty on shape discrimination thresholds, Green *et al.* (2017, 2018a,b,c) found that observers' were more sensitive at using changes in curvature to discriminate between shapes when the phase of shapes (i.e., angular position) varied from trial to trial. This findings suggest that spatial uncertainty modulates how curvatures are processed along the contour of objects (Green *et al.*, 2017, 2018a,b,c), but does not reveal how spatial uncertainty might impact visual sensitivity to curvatures along isolated angular subtenses of the visual field.

The purpose of the current study, therefore, is to test whether human observers are sensitive to positive deformations of curvature across polar angles. We make use of a novel stimulus class, which we refer to as Difference of Gaussian contours, that grant us greater control over the parameters that define regions of deformation along a closed contour stimulus. We believe that our results are the first to demonstrate that observers are indeed uniformly sensitive to positive deformations of curvature at all polar angles tested, both in the presence of absence of spatial uncertainty, and across different angular frequencies.

4.3 General Methods

4.3.1 Participants

Three experienced psychophysical observers participated in the experiment. Two observers were naïve to the purpose of the study (LUC, KAT), and one was an author (MS). The mean age of the observers was 24.7 ($SD = 2.52$, range: 22-27) and all observers had normal or corrected-to-normal visual acuity. Experimental protocols were approved by the McMaster University Research Ethics Board, and informed consent was obtained from each observer prior to the start of the experiment.

4.3.2 Apparatus

An APPLE G4 2.66 GHz Quad-Core Intel Xeon computer generated and displayed stimuli using Matlab 10.7.0 (Mathworks Inc., Natick, MA) and Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) on a Sony Model GDM-F520 monitor, with a pixel resolution 1024 × 768 (62.5 pixels/deg) and a refresh rate of 100 Hz. The display was the only light source in the room during testing, and had a mean luminance of 69.2 cd/m². The stimuli were viewed binocularly through natural pupils. A chin rest was used to ensure a viewing distance of 131 cm was maintained throughout the duration of the experiment.

4.3.3 Stimuli: Difference of Gaussian contours

Our stimuli were Difference of Gaussian (DoG) contours, which are defined as:

$$r(\theta) = \bar{r}(1 + DoG(\theta)) \qquad (4.1)$$

126

where $\bar{r}$ is the mean radius of the base circle, and the DoG corresponds to the deformed segment along the circumference of a base circle defined by:

$$
\begin{aligned}
DoG(\theta) &= G_{\sigma_1}(\theta; \theta_c; \sigma_1) - G_{\sigma_2}(\theta; \theta_c; \sigma_2) \\
&= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sigma_1} e^{-(\theta-\theta_c)/2\sigma_1^2} - \frac{1}{\sigma_2} e^{-(\theta-\theta_c)/2\sigma_2^2} \right)
\end{aligned}
\tag{4.2}
$$

where G_{σ_1} and G_{σ_2} are Gaussian probability density functions. Parameter θ_c and variance (σ^2) of each normal distribution determine the angular frequency, amplitude, and location of the deformed segment along the circumference of a base circle. The DoG portion of the contour smoothly transitions into a base circle with a radius equal to $\bar{r}$. To ease the process of choosing parameters for the Gaussian functions that compose the DoG function, we fit Equation 4.3 to one positive cycle of angular frequency, where the equation for angular frequency is:

$$
r(\theta) = \bar{r}(1 + A sin(\omega\theta + \phi))
\tag{4.3}
$$

where $\bar{r}$ is the base radius of the circular contour, A is the amplitude, ω is the frequency, and ϕ is the phase (also see Eq. 1 in Wilkinson *et al.*, 1998).

The main benefit of using DoG contours, as opposed to radial frequency (RF) patterns described by Wilkinson *et al.* (1998), is that a single region of the close contour can be modulated by frequencies that are non-integer values. This feature makes it easier to investigate the perception of deformed regions that vary along a continuum of angular frequency.

Examples of DoG contours used in the present study, fit to angular frequencies of 3 and 6 cycles of modulation per circumference, are shown in Figure 4.1. In the current experiments, we measured curvature detection thresholds with stimuli in which the location of the DoG region varied across 24 polar angles (in 15° increments).

The radial luminance profile of each contour was defined by a fourth derivative Gaussian (D4; see Wilkinson *et al.*, 1998) with a luminance contrast set to 99% and a peak spatial frequency of 8 cpd. All DoG contours had a base radius ($\bar{r}$) of 1.14°.

4.3.4 Psychophysical Procedure

Thresholds were obtained using a two-interval forced choice (2IFC) task. On each trial, observers were presented with a modulated DoG contour in one stimulus interval and a circular (i.e., non-modulated) contour in the other interval. Observers were asked to judge whether the first or second interval contained the modulated stimulus. A 60 s light adaptation period preceded the start of every testing session to ensure observers eyes were adapted to the average luminance of the display. Each trial began with a small, high-contrast fixation dot drawn in the middle of the display. Observers initiated trials by pressing the spacebar, after which the fixation dot flickered at 10 Hz for 200 ms. The fixation point was extinguished and, after a 300 ms delay, was followed by the two 150 ms stimulus intervals separated by a 500 ms ISI. Across stimulus intervals and trials, the central position of each contour was jittered randomly 0.17° around a circular region originating from the center of the screen. Observers received 50 practice trials prior to the start of their first experimental session, and 10 practice trials at the start of subsequent testing sessions. Auditory feedback that informed the observers as to the correctness of their response on every trial throughout the
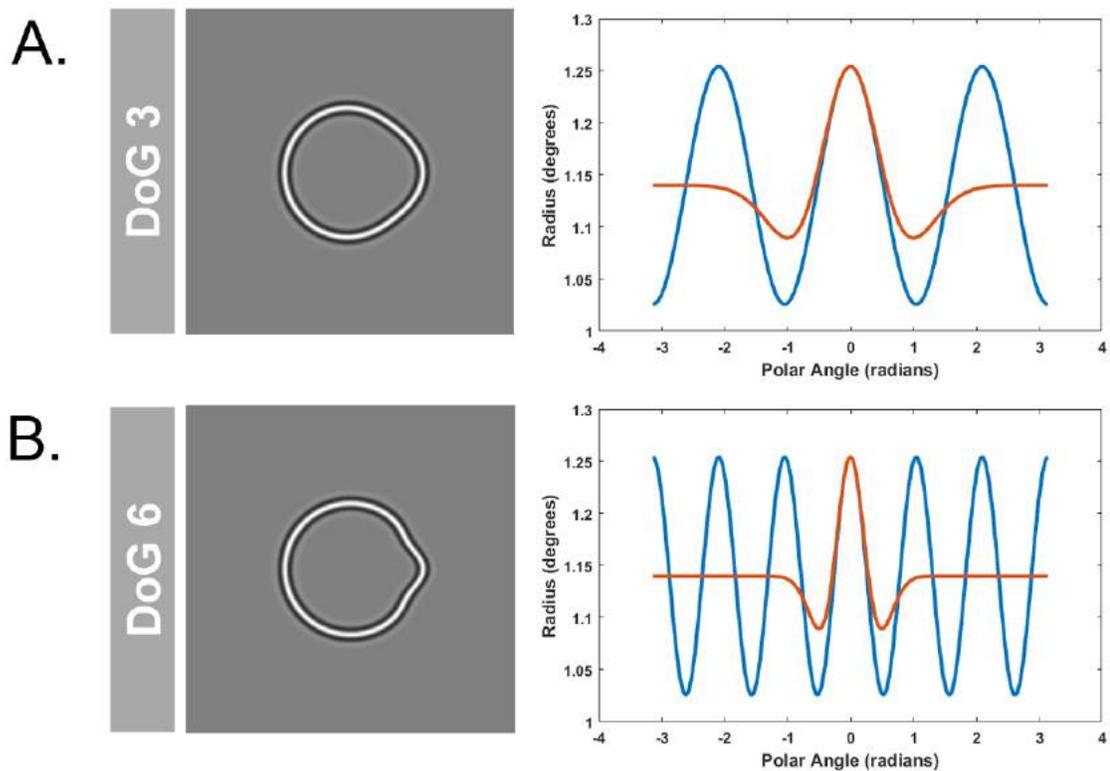
Figure 4.1: Example of DoG contours used in this study. The amplitude of modulated DOG contours shown are set to 10% of the radius of the base circle for illustrative purposes. A) DoG contour fit to an angular frequency of 3 cycle/2π. B) DoG contour fit to a angular frequency of 6 cycle/2π. In the experiments, the luminance profile of the contour was a D4 pattern, with a peak spatial frequency of 8 cpd and a luminance contrast of 99%.

experiment.

In the fixed angle condition, the polar angle of the DoG function was fixed within a single testing session, but varied across sessions. In the random angle condition, the polar angle of the DoG function varied randomly across 24 angles within each testing session. Within each session, the amplitude of the DoG function was adjusted across trials with 24 independent, 2-down, 1-up staircases (one for each angle) that converged on 71% correct performance. Each staircase terminated after 16 reversals. Thresholds were estimated by averaging the modulation amplitude of the last eight reversals from each staircase. Observers completed two sessions per DoG contour and angle uncertainty condition, for a total of eight testing sessions.

4.3.5 Data analysis

All analyses were performed using the statistical computing software R (R Core Team, 2017). Mixed linear models were used to analyze thresholds using the lme4 package (Bates *et al.*, 2015). The Kenward-Roger method (Kenward and Roger, 1997) was used to approximate the degrees of freedom for each mixed model, as this method better approximates F-distributions for linear mixed models (Judd *et al.*, 2012). For brevity, we report only the F tests from the linear mixed-effects regression analyses (i.e., the analysis of variance of Type III sums of squares with Kenward-Roger approximation for degrees of freedom). To increase the power of our statistical analyses, we performed a more focused comparison of curvature detection thresholds across four visual quadrants (see Figure 4.3). Each quadrant was composed of detection thresholds collected across 6 polar angles that split the visual field both horizontally (top and bottom), and vertically (left and right).

4.3.6   Results

Results for each observer are shown in Figure 4.2. Thresholds were lower for smaller curvature deformations (DoG 6) than larger deformations (DoG 3). Furthermore, curvature detection thresholds were lower when the polar angle of peak curvature was fixed within a block of trials, compared to when the angle varied randomly across trials. Finally, in all three observers sensitivity to curvature was approximately uniform across all polar angles tested. Aggregate results with polar angles grouped according to visual quadrant are shown in Figure 4.3. The data were analyzed with a mixed linear model that included Visual Quadrant, DoG angular frequency, and Angle Uncertainty as fixed effects, and Observer and Session as random effects. The ANOVA revealed a significant effect of DoG angular frequency ($F_{1,555} = 300.25$, $p < .0001$) and Angle Uncertainty ($F_{1,555} = 19.34$, $p < .0001$), with no other terms reaching significance ($p > .35$ in each case).

To estimate an oblique effect, we used a linear contrast to compare thresholds measured with oblique (45°, 135°, 225°, 315°) and cardinal (0°, 90°, 180°, 270°) orientations. The linear contrast essentially computes the difference between the average thresholds for the oblique and cardinal orientations. The average thresholds are shown in Figure 4.4 A: for observers KAT and LUC, thresholds were 22% and 28% higher for oblique orientations than cardinal orientations, whereas observer MS exhibited nearly identical thresholds for the two sets of orientations. On average, thresholds were 14.5% higher for oblique orientations, a difference that was not statistically significant ($t(2) = 1.61$, $p = 0.25$).
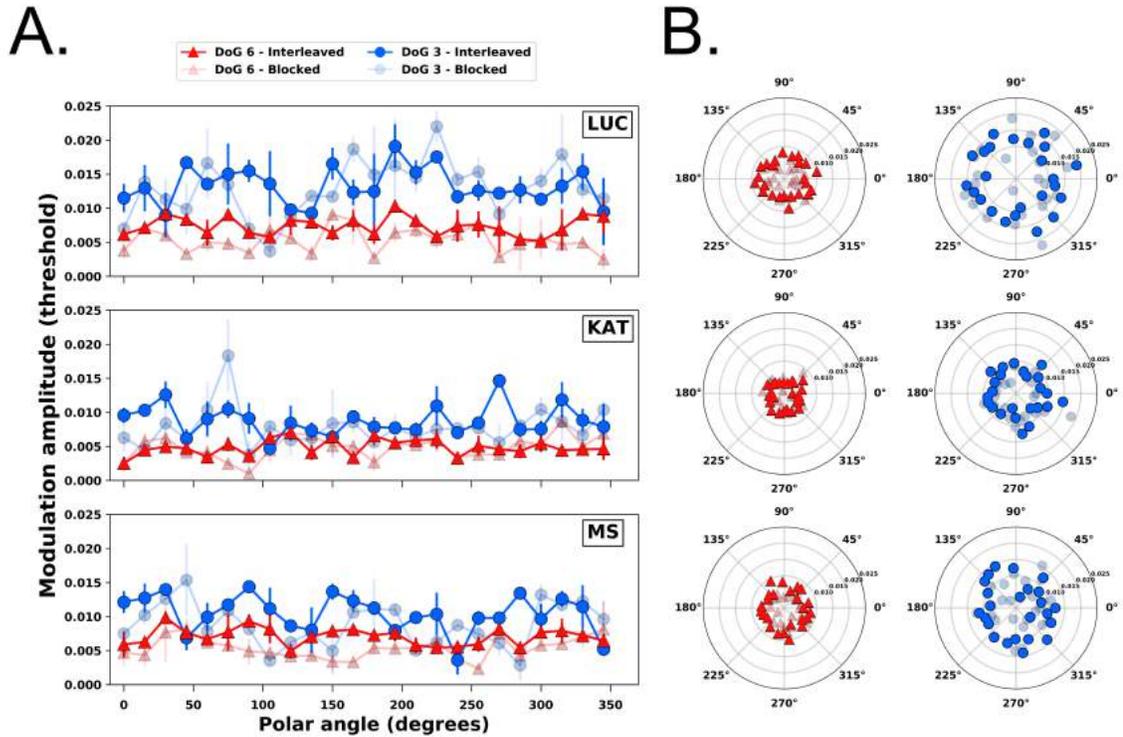
Figure 4.2: A) Curvature detection thresholds for three observers plotted as a function of polar angle of peak positive curvature. The angular frequency of the DOG contour is represented by symbol colour and shape, while transparency of lines distinguish between manipulations of polar angle uncertainty. Error bars represent ±1 SEM. B) Polar plots of mean detection thresholds for each observer for all conditions.
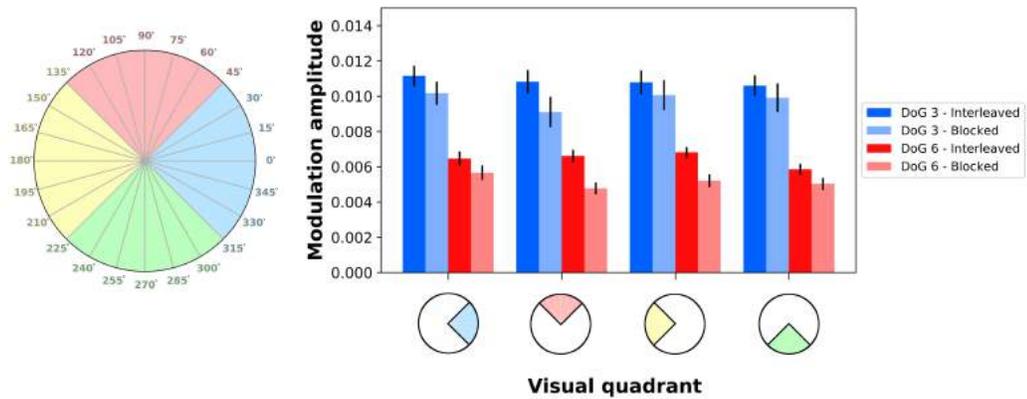
Figure 4.3: Average curvature detection thresholds grouped according to visual quadrant. Each visual quadrant contains thresholds from 6 separate polar angle conditions. Dark bars represent conditions where polar angle varied from trial to trial, with lighter bars representing conditions where polar angle location was fixed within a block of trials. Angular frequency is represented by the colour of each bar (DoG3 - blue; DoG6 - red), and errors bars represent ± 1 SEM.
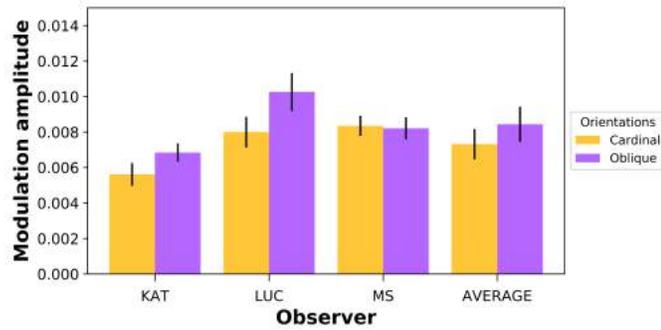
(a) Experiment 1
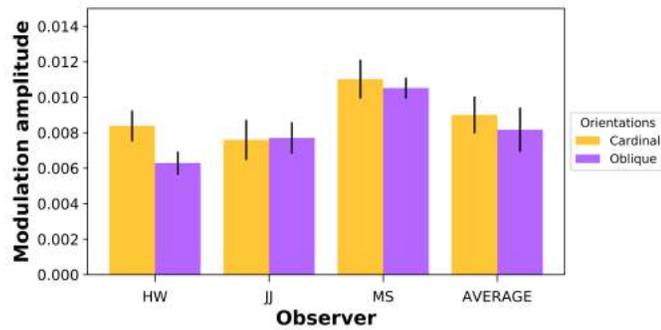
(b) Experiment 2

Figure 4.4: Thresholds measured at oblique and cardinal orientations in Experiments 1 and 2. Average thresholds are shown individual observers as well as averaged across observers. Error bars are $\pm 1\,\mathrm{SEM}$.

4.3.7 Discussion

Results from Experiment 1 demonstrate that observers are approximately uniformly sensitive to curvature deformations as a function of polar angle, regardless of angular frequency or whether spatial uncertainty is introduced in the location of deformation. However, curvature detection thresholds are lower for the higher angular frequency and when polar angle was constant within a block of trials. This last result suggests that curvature detection thresholds are affected by spatial uncertainty about the polar angle of the deformation.

One possible explanation for the uniform sensitivity to curvature across polar angles is that observers performed the task by looking for a circle, not the deformed target. By identifying which interval contained the circle, observers could attenuate the effects of polar angle and spatial uncertainty on curvature detection thresholds, as there would be no need to direct attention towards regions where curvature changes occur. However, if observers were performing the task by identifying the interval that contained the circle, then they would be unable to identify the polar angle of the curvature deformation. Experiment 2 examined this possibility by measuring curvature detection thresholds under conditions of spatial uncertainty using a dual-judgment task, in which observers were asked to identify both the interval containing the curvature deformation as well as its orientation.

4.4 Experiment 2

4.4.1 Participants

Two new naïve observers (HW and JJ) participated in Experiment 2, along with the author (MS) that had participated in Experiment 1. The mean age of the observers was 22.0 ($SD = 4.35$, range: 19-27) and all observers had normal or corrected-to-normal visual acuity. Experimental protocols were approved by the McMaster University Research Ethics Board, and informed consent was obtained from each participant prior to the start of the experiment.

4.4.2 Apparatus, Stimuli, & Procedure

The apparatus, stimuli, and psychophysical procedure were identical to those described for Experiment 1 except for the following minor alterations to the procedure. First, observers were tested only in random angle conditions where the location of the deformation varied across trials. Second, the location of the deformation was limited to 8 polar angles (i.e,. 0 to 315° in 45° increments). Finally, observers were asked to identify both the interval and location of the deformity using the number-pad on a computer keyboard. Observers completed two sessions per DoG condition, for a total of four testing sessions.

4.4.3 Data analysis

Data were analyzed using mixed-linear models as in Experiment 1, except that polar angles were not grouped into visual quadrants. Analyses of localization judgements were restricted to the responses on trials occurring between reversals 8 and

16 on each staircase to ensure that judgements were made for stimuli that included curvature deformations that were near detection threshold.

4.4.4   Results

Curvature detection thresholds for Experiment 2 are shown in Figure 4.5. As was found in Experiment 1, curvature detection thresholds were approximately constant across polar angle, and were lower for the higher angular frequency (i.e, 6 cycles/2π vs. 3 cycles/2π), sine wave. These observations were tested using a mixed linear model with Polar Angle and DoG angular frequency as fixed effects, and Observer and Session as random effects. The ANOVA revealed a significant effect of DoG angular frequency ($F_{1,79.45} = 41.60$, $p < .0001$), with no other terms approaching significance ($p > .47$ in each case).

As in Experiment 1, we performed a linear comparison of detection thresholds measured with oblique versus cardinal orientations. The average thresholds at the two sets of orientations are shown in Figure 4.4 B: for observers HW and MS, thresholds were 25% and 5% *lower* for oblique orientations than cardinal orientations, whereas observer JJ exhibited nearly identical thresholds for the two sets of orientations. On average, thresholds were 9% lower for oblique orientations, a difference that was not statistically significant ($t(2) = 0.85$, $p = 0.48$).

The proportion of correct angle identification judgements are shown in Figure 4.6a. For both DoG angular frequencies, and across all polar angles, observers were able to identify the orientation of the deformity at accuracies that were well-above chance. The proportions of correct angle identification responses were analyzed with a mixed-linear model that included two fixed effects (DoG angular frequency and Polar
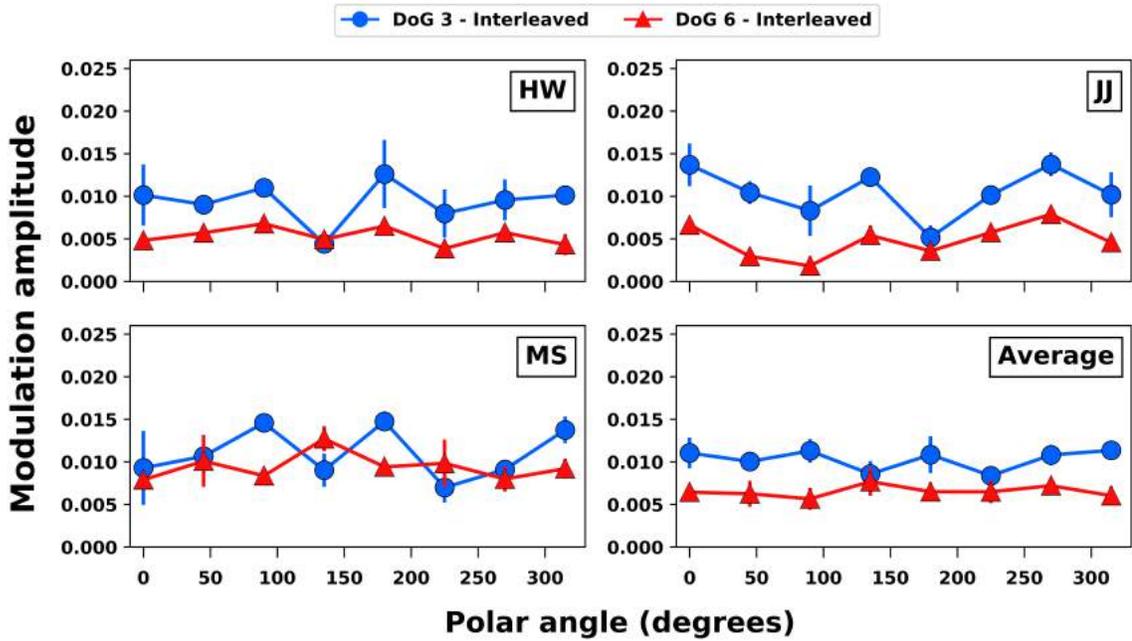
Figure 4.5: Curvature detection thresholds measured in Experiment 2 for three observers plotted as a function of polar angle of peak positive curvature. The angular frequency of the DoG contour is coded for by line colour and shape, with error bars representing ±1 SEM.
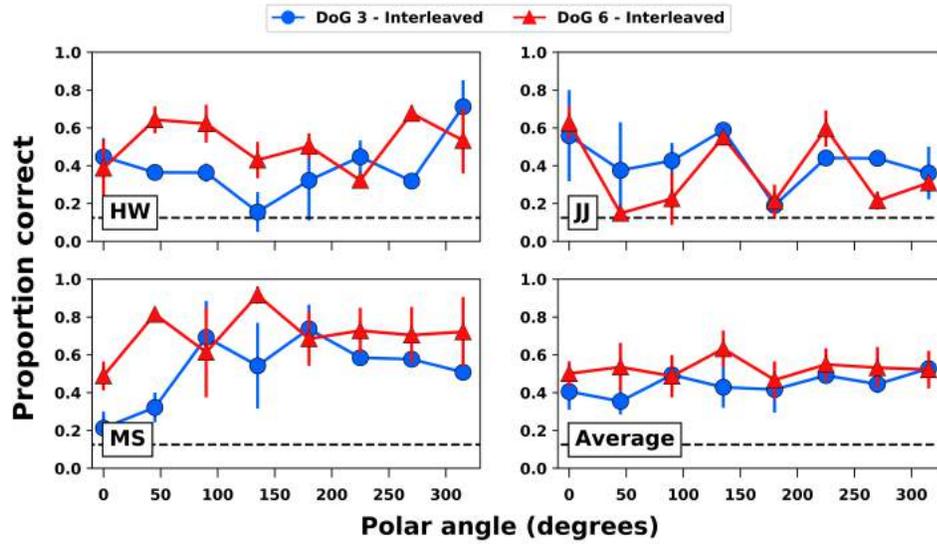
(a)
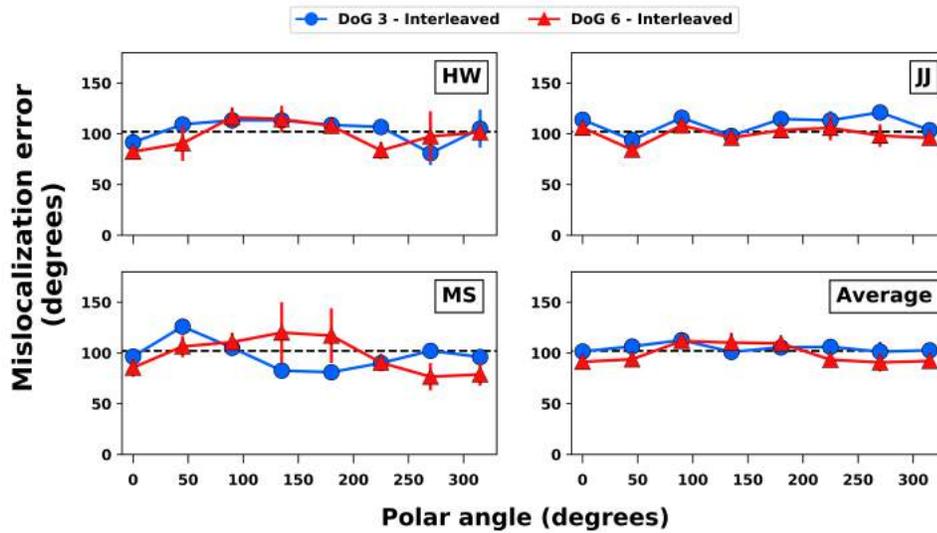
(b)

Figure 4.6: Polar angle identification results from Experiment 2 on trials in which the curvature was near detection threshold. a) Mean accuracy of angle identification judgements for each angular frequency and each polar angle. The dashed lines indicate chance performance. b) Mean absolute error in angle identification responses (in degrees of polar angle) plotted as a function of DoG angular frequency and polar angle. Dashed lines represent the predicted error based on guessing. Error bars represent ±1 SEM.

Angle), and two random effects (Observer and Session). Results from the analysis were consistent with the observations described above, in that no effects reached significance ($p > .20$ in each case).

The magnitude of angle identification error is plotted as a function of angular frequency and polar angle in Figure 4.6b. Error magnitude did not differ markedly across angular frequencies or polar angles, and an analysis with a mixed effect model that included the same factors as the one used to analyze identification accuracy revealed no significant effects ($p > .11$ in each case).

We next considered whether the identification errors were consistent with the hypothesis that observers were guessing. If all identification errors were caused by guessing, and if all directions were equally likely to be guessed, then the average error ought to be 102 deg. Inspection of Figure 4.6 b suggests that the errors were very close to the predicted value for guessing: the mean errors for the three observers were 101.35, 104.44, and 97.88, and the grand mean was 101.2. Hence, the observed errors were consistent with the hypothesis that most, if not all, errors were guesses.

Relating detection & identification

Finally, we examined the association between curvature detection and angle identification judgements. Detection and identification judgements on near-threshold trials (i.e., trials between staircase reversals 8 and 16) are shown in Table 4.1. For each subject, the probability of correctly identifying the polar angle was much greater when the detection judgement was correct, and the probability of correctly detecting the curvature deformation was much greater on trials on which the identification judgment was correct. These observations were supported by chi-square tests that rejected

Table 4.1: Contingency tables for detection and identification judgements for near-threshold trials (between staircase reversals 8 & 16) in Experiment 2.

Observer HW		Identification	
		correct	incorrect
Detection	correct	337	261
	incorrect	65	213

Observer JJ		Identification	
		correct	incorrect
Detection	correct	308	293
	incorrect	53	259

Observer MS		Identification	
		correct	incorrect
Detection	correct	366	93
	incorrect	84	106

the hypothesis of no association between detection and identification ($\chi^2_{(1)} \geq 78.1$, $p < .0001$ for each observer).

The association between detection and identification is illustrated in a different manner, and a across a wider range of curvature amplitudes, in Figure 4.7. Figure 4.7a plots the proportion of correct detection responses (after collapsing across conditions, sessions, and observers) as a function of log-transformed curvature amplitude, separately for trials on which the identification response was correct and incorrect. For trials on which the identification judgements were correct, detection accuracy increased with increasing curvature amplitude, and the data were well-fit by a Weibull function. On the other hand, detection accuracy on trials where the identification judgement was incorrect hovered slightly above chance levels and did not change noticeably with increasing curvature amplitude. As a result, the psychometric function fit to the data from identification-correct trials is shifted to lower amplitudes and is noticeably shallower than the function fit to all of the data. Figure 4.7b shows identification accuracy plotted against modulation amplitude, separately for trials on which detection responses were correct and incorrect. As was found with detection accuracy, identification accuracy was much lower on trials when detection responses were incorrect; however, unlike what was found for detection, there is some suggestion that identification accuracy on detection-incorrect trials increased slightly with increasing deformation amplitude. Consequently, and unlike what was found for detection accuracy, the psychometric function fit to identification accuracy for detection-incorrect trials is similar to the function fit to all of the data.

In summary, our results are consistent with the idea that detection and identification judgements were not independent: correct detection was associated with

(a) Detection accuracy.

(b) Identification accuracy.
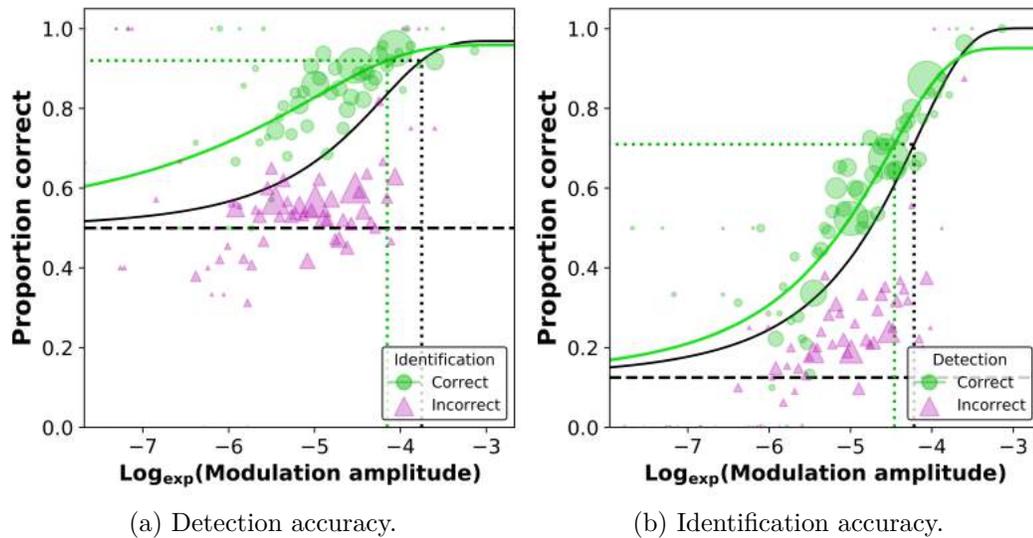
Figure 4.7: a) Detection accuracy plotted against log-transformed modulation amplitude, separately for trials on which identification was correct or incorrect. b) Identification accuracy plotted separately for trials on which detection was correct or incorrect. In both figures, the data were collapsed across conditions and observers, and the number of trials for each point is represented by symbol size. The green lines are the best-fitting (weighted least-squares) Weibull function fit to data from correct (a) identification or (b) detection responses; the black lines are the Weibull functions fit to all of the data. The horizontal dashed lines represent chance performance, and the dotted lines indicate points on the psychometric functions where d' equals 2, which corresponds to 92% and 71% accuracy in the detection and identification tasks, respectively.

correct identification, and *vice versa.* This association was not an artifact caused by detection and identification accuracy both being positively correlated with deformation amplitude, because it was found even when we considered only trials near detection threshold (Table 4.1). When we examined the association between the two judgements across a wider range of stimulus levels, we found that detection and identification accuracy were both near chance and only weakly associated with deformation amplitude when responses on the other task were incorrect (Figure 4.5). This second results implies that observers had very poor information about polar angle when the curvature deformation was not detected, and very poor information about the stimulus interval containing the target when its polar angle was misidentified.

4.5   General Discussion

The current study used a novel stimulus class we call DoG contours to examine the sensitivity of human observers to local curvature deformations along a closed circular contour as a function of polar angle. We found that thresholds were higher when observers were uncertain about polar angle, but that observers were approximately equally sensitive to curvature deformations across all polar angles tested (0 - 345°). We found no evidence for an oblique effect for curvature detection. Furthermore, observers accurately identified the angle of deformities when the magnitude of the deformation was near detection threshold and the location of the deformity was uncertain. Together, these results suggest that observers are approximately uniformly sensitive to curvature deformations across all polar angles.

4.5.1 Angular frequency

Consistent with previous findings, we found that observers were more sensitive to curvature segments of higher angular frequency (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). Data from studies that measured detection thresholds for single cycles of sinusoidal deformation along closed contours suggest that performance is constrained by mechanisms that encode local contour shape rather than the global shape of the enclosed contour (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). For example, Schmidtmann *et al.* (2012) found that detection thresholds decreased as the number of sinusoidal deformation cycles increased, and that thresholds obtained with fewer than three cycles were well fit by probability summation models that combined responses across multiple, independent local mechanisms. However, Schmidtmann and Kingdom (2017) demonstrated that curvature detection thresholds also decrease for sinusiodally modulated lines of increasing frequency similar to that of angular frequency for closed contour stimuli. Therefore, increased sensitivity to curvatures of high frequency may reflect a general mechanism that can encode higher frequency curvatures better, whether the modulation is applied to a circle or a line stimulus (Schmidtmann and Kingdom, 2017).

Regardless of how curvatures are summed along contours, local curvature detectors are thought to be based on analyses of local orientation that are performed in V1 or V2 (Blakemore and Over, 1974; Hubel and Wiesel, 1965) and form the input to shape-encoding mechanisms that operate over larger spatial scales (Felleman and Van, 1991). Computational models that incorporate such a hierarchical arrangement of visual analyses have been used to predict the population responses of V4 cells (Pasupathy and Connor, 2001, 2002; Rodríguez-Sánchez and Tsotsos, 2012), which

are thought to integrate local curvature. Our data are consistent with these types of models, as observers were more sensitive at both detecting and localizing curvatures that were restricted to smaller fractions along a closed contour.

4.5.2   Uniform sensitivity to curvature & localizability

The current experiments demonstrate that observers are approximately equally sensitive to local curvature across all polar angles. Neurophysiological evidence from macaques suggest this uniform sensitivity may arise from the uniform sampling of polar angle by neurons sensitive to positive curvature deformations in V4 (Pasupathy and Connor, 2001, 2002). However, it is unclear whether responses from V4 neurons need to be considered in characterizing performance on this task. Another possible explanation is that observers based their judgements on the responses of V1 and/or V2 neurons, which is consistent with psychophysical evidence from single cycle studies of RF patterns (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). For example, hypercomplex cells in primary visual cortex have been shown to be responsive to contour curvature (Dobbins *et al.*, 1987, 1989; Hubel and Wiesel, 1965), and therefore could serve as a mechanism for detecting local curvature, without the need to monitor responses from cells at higher visual areas.

Data from neuroimaging studies have shown that neurons in human primary visual cortex oriented along cardinal meridians are over represented and exhibit stronger responses to stimuli relative to neurons tuned to oblique angles (Furmanski and Engel, 2000; Li *et al.*, 2003). Interestingly, these differences between cardinal and oblique orientations appear to be less conspicuous in visual cortical areas beyond V1 (i.e., V2, V3) (Furmanski and Engel, 2000; Levitt *et al.*, 1994). Hence, differences between V1

neurons tuned to cardinal and oblique orientations are thought to underlie oblique effects, which are especially apparent for stimuli located near the fovea (Berkley *et al.*, 1975; Vandenbussche *et al.*, 1986), like the stimuli used in our experiments. Therefore, if neurons in primary visual cortex were being used to encode curvature deformations, sensitivity to curvature deformations should be better along cardinal than oblique orientations. Contrary to this prediction, we found that sensitivity to curvature was approximately constant across all polar angles.

Localization responses across observers from Experiment 2 also suggest that once changes in curvature are reliably detected (i.e., reach threshold), curvatures can be localized well above chance, even when the location of the deformity varies from trial to trial. Although this finding alone cannot be used to differentiate between models described above in detecting curvature, it does suggest that the mechanism(s) used to detect curvature are also sensitive to the location at which deformity occurred. Therefore, location of changes in curvature is an additional requirement that should be captured by a model aimed at describing how the visual system represents curvatures along closed contours. For example, the population response of V4 neurons as modelled by Pasupathy and Connor (2002) represents closed contour shapes along two dimensions, curvature and polar angle, that would allow an observer to both locate and detect curvature along a shape using the same mechanism. In contrast, models that represent curvatures as the weighted sum of periodic angular frequencies, such the later stages in select models of human shape perception (Poirier and Wilson, 2006), would have trouble producing neural codes useful in localizing curvature deformities. This is because location information would be lost after a Fourier-like decomposition of contour curvature. Future work should further investigate whether,

or when location specific information is lost in representing curvatures along shapes.

4.5.3 Spatial uncertainty

The results of our study also demonstrate that spatial uncertainty, varied by fixing or randomizing the location of peak deformation along a contour, modulates detection thresholds. These results are consistent with previous psychophysical studies whereby attenuating attentional effects through the removal of spatial uncertainty led to improved performance at detecting modulated contour fragments along circular shapes (Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). Attention has been shown to increase the discriminability of select types of spatial information, especially when focused towards a well localized region of space (Lee *et al.*, 1997). One possibility is that constricting visual attention to a well localized portion of the visual field is associated with an increase in processing power of information that falls within that region (Eriksen and Yeh, 1985). Studies of spatial selective attention in macaques have demonstrated an increase in the spontaneous firing rates of neurons when attention is directed inside the receptive field of a neuron, even in the absence of visual information (Luck *et al.*, 1997). Neurons within area V4, a candidate area for curvature processing, have been shown to increase their response to attended targets by as much as 51% (Mitchell *et al.*, 2007; Reynolds *et al.*, 2000), and also increase their neural synchrony of firing rates to attended locations (Taylor *et al.*, 2005). Therefore, attention may be acting as a mechanisms that controls the gain of signals representing curvature information. This also helps to explain why observers are able to detect curvature deformities once well localized, such as the findings in Experiment 2, as attended locations receive a boost in signal strength. Future work should aim

to elucidate the role attention has at modulating curvature information along closed contours at different levels of processing along the ventral visual stream.

4.6 Conclusion

In summary, the aim of our study was to test the hypothesis that observers are uniformly sensitive to positive curvature deformations across all polar angles. Using a novel stimulus class we call Difference of Gaussian (DOG) contours, we have shown that observers are indeed uniformly sensitive at detecting positive deformations of curvature, regardless of spatial uncertainty or angular frequency of deformed parts. Our study is the first to demonstrate that human sensitivity to curvature deformations is well preserved across a polar coordinate system.

4.7   References

Anzai, A., Peng, X., and Van Essen, D. C. (2007). Neurons in monkey visual area V2 encode combinations of orientations. *Nature Neuroscience*, **10**(10), 1313–1321.

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: the "oblique effect" in man and animals. *Psychological Bulletin*, **78**(4), 266–278.

Baldwin, A. S., Schmidtmann, G., Kingdom, F. A. A., and Hess, R. F. (2016). Rejecting probability summation for radial frequency patterns, not so Quick! *Vision Research*, **122**, 124–134.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48.

Bell, J. and Badcock, D. R. (2009). Narrow-band radial frequency shape channels revealed by sub-threshold summation. *Vision Research*, **49**(8), 843–50.

Bell, J., Wilkinson, F., Wilson, H. R., Loffler, G., and Badcock, D. R. (2009). Radial frequency adaptation reveals interacting contour shape channels. *Vision Research*, **49**(18), 2306–17.

Berkley, M. A., Kitterle, F., and Watkins, D. W. (1975). Grating visibility as a function of orientation and retinal eccentricity. *Vision Research*, **15**(2), 239–244.

Blakemore, C. and Over, R. (1974). Curvature Detectors in Human Vision? *Perception*, **3**(1), 3–7.

Bonds, A. B. (1982). An "oblique effect" in the visual evoked potential of the cat. *Experimental Brain Research*, **46**(1), 151–154.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433–36.

Campbell, F. W., Kulikowski, J. J., and Levinson, J. (1966). The effect of orientation on the visual resolution of gratings. *The Journal of Physiology*, **187**(2), 427–436.

Dakin, S. C., Mareschal, I., and Bex, P. J. (2005). An oblique effect for local motion: Psychophysics and natural movie statistics. *Journal of Vision*, **5**(10), 9–9.

Dobbins, A., Zucker, S. W., and Cynader, M. S. (1987). Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature*, **329**(6138), 438–441.

Dobbins, A., Zucker, S. W., and Cynader, M. S. (1989). Endstopping and curvature. *Vision Research*, **29**(10), 1371–1387.

Eriksen, C. W. and Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of experimental psychology. Human perception and performance*, **11**(5), 583–597.

Felleman, D. J. and Van, D. E. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, N.Y. : 1991)*, **1**(1), 1–47.

Furmanski, C. S. and Engel, S. A. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience*, **3**(6), 535–536.

Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W., and Essen, D. C. V. (1996). Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*, **76**(4), 2718–2739.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2017). Global processing of random-phase radial frequency patterns but not modulated lines. *Journal of Vision*, **17**(9), 18–18.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018a). Convergent evidence for global processing of shape. *Journal of Vision*, **18**(7), 7–7.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018b). The effect of spatiotemporal displacement on the integration of shape information. *Journal of Vision*, **18**(5), 4–4.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018c). Integration of shape information occurs around closed contours but not across them. *Journal of Vision*, **18**(5), 6–6.

Gros, B. L., Blake, R., and Hiris, E. (1998). Anisotropies in visual motion perception: a fresh look. *JOSA A*, **15**(8), 2003–2011.

Heeley, D. W., Buchanan-Smith, H. M., Cromwell, J. A., and Wright, J. S. (1997). The oblique effect in orientation acuity. *Vision Research*, **37**(2), 235–242.

Hegdé, J. and Essen, D. C. V. (2000). Selectivity for Complex Shapes in Primate Visual Area V2. *Journal of Neuroscience*, **20**(5), RC61–RC61.

Hess, R. F., Wang, Y. Z., and Dakin, S. C. (1999). Are judgements of circularity local or global? *Vision Res*, **39**(26), 4354–60.

Hubel, D. H. and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, **148**(3), 574–591.

Hubel, D. H. and Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, **28**(2), 229–289.

Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, **195**(1), 215–243.

Hubel, D. H., Wiesel, T. N., and Stryker, M. P. (1978). Anatomical demonstration of orientation columns in macaque monkey. *The Journal of Comparative Neurology*, **177**(3), 361–379.

Ito, M. and Komatsu, H. (2004). Representation of Angles Embedded within Contour Stimuli in Area V2 of Macaque Monkeys. *Journal of Neuroscience*, **24**(13), 3313–3324.

Jeffrey, B. G., Wang, Y. Z., and Birch, E. E. (2002). Circular contour frequency in shape discrimination. *Vision Research*, **42**(25), 2773–9.

Judd, C. M., Westfall, J., and Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, **103**(1), 54–69.

Kenward, M. G. and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, **53**(3), 983–997.

Lee, D. K., Koch, C., and Braun, J. (1997). Spatial vision thresholds in the near absence of attention. *Vision Research*, **37**(17), 2409–2418.

Levitt, J. B., Kiper, D. C., and Movshon, J. A. (1994). Receptive fields and functional architecture of macaque V2. *Journal of Neurophysiology*, **71**(6), 2517–2542.

Li, B., Peterson, M. R., and Freeman, R. D. (2003). Oblique Effect: A Neural Basis in the Visual Cortex. *Journal of Neurophysiology*, **90**(1), 204–217.

Loffler, G., Wilson, H. R., and Wilkinson, F. (2003). Local and global contributions to shape discrimination. *Vision Res*, **43**(5), 519–30.

Luck, S. J., Chelazzi, L., Hillyard, S. A., and Desimone, R. (1997). Neural Mechanisms of Spatial Selective Attention in Areas V1, V2, and V4 of Macaque Visual Cortex. *Journal of Neurophysiology*, **77**(1), 24–42.

Mansfield, R. J. W. (1974). Neural Basis of Orientation Perception in Primate Vision. *Science*, **186**(4169), 1133–1135.

Mitchell, J. F., Sundberg, K. A., and Reynolds, J. H. (2007). Differential Attention-Dependent Response Modulation across Cell Classes in Macaque Visual Area V4. *Neuron*, **55**(1), 131–141.

Pasupathy, A. and Connor, C. E. (1999). Responses to contour features in macaque area v4. *Journal of Neurophysiology*, **82**(5), 2490–2502.

Pasupathy, A. and Connor, C. E. (2001). Shape representation in area v4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, **86**(5), 2505–2519.

Pasupathy, A. and Connor, C. E. (2002). Population coding of shape in area v4. *Nature neuroscience*, **5**(12), 1332—1338.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysocs: transforming numbers into movies. *Spatial Vision*, **10**, 437–42.

Poirier, F. J. and Wilson, H. R. (2006). A biologically plausible model of human radial frequency perception. *Vision Research*, **46**(15), 2443 – 2455.

R Core Team (2017). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Reynolds, J. H., Pasternak, T., and Desimone, R. (2000). Attention Increases Sensitivity of V4 Neurons. *Neuron*, **26**(3), 703–714.

Riesenhuber, M. and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, **2**, 1019–1025.

Rodríguez-Sánchez, A. J. and Tsotsos, J. K. (2012). The Roles of Endstopped and Curvature Tuned Computations in a Hierarchical Representation of 2d Shape. *PLOS ONE*, **7**(8), e42058.

Schmidtmann, G. and Kingdom, F. A. A. (2017). Nothing more than a pair of curvatures: A common mechanism for the detection of both radial and non-radial frequency patterns. *Vision Research*, **134**, 18–25.

Schmidtmann, G., Kennedy, G. J., Orbach, H. S., and Loffler, G. (2012). Non-linear global pooling in the discrimination of circular and non-circular shapes. *Vision Research*, **62**, 44–56.

Schmidtmann, G., Logan, A. J., Kennedy, G. J., Gordon, G. E., and Loffler, G. (2015). Distinct lower visual field preference for object shape. *Journal of Vision*, **15**(5), 18–18.

Tanaka, K. (2003). Columns for Complex Visual Object Features in the Inferotemporal Cortex: Clustering of Cells with Similar but Slightly Different Stimulus Selectivities. *Cerebral Cortex*, **13**(1), 90–99.

Taylor, K., Mandon, S., Freiwald, W. A., and Kreiter, A. K. (2005). Coherent Oscillatory Activity in Monkey Area V4 Predicts Successful Allocation of Attention. *Cerebral Cortex*, **15**(9), 1424–1437.

Teller, D. Y., Morse, R., Borton, R., and Regal, D. (1974). Visual acuity for vertical and diagonal gratings in human infants. *Vision Research*, **14**(12), 1433–1439.

Vandenbussche, E., Vogels, R., and Orban, G. A. (1986). Human orientation discrimination: changes with eccentricity in normal and amblyopic vision. *Investigative Ophthalmology & Visual Science*, **27**(2), 237–245.

Wang, G., Ding, S., and Yunokuchi, K. (2003). Representation of Cardinal Contour Overlaps Less With Representation of Nearby Angles in Cat Visual Cortex. *Journal of Neurophysiology*, **90**(6), 3912–3920.

Westheimer, G. (2003). Meridional anisotropy in visual processing: implications for the neural site of the oblique effect. *Vision Research*, **43**(22), 2281–2289.

Wilkinson, F., Wilson, H. R., and Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Res*, **38**(22), 3555–68.

Xu, X., Collins, C. E., Khaytin, I., Kaas, J. H., and Casagrande, V. A. (2006). Unequal representation of cardinal vs. oblique orientations in the middle temporal visual area. *Proceedings of the National Academy of Sciences*, **103**(46), 17490–17495.

Chapter 5

Evaluating spatio-temporal interactions between shapes

5.1 Abstract

Spatio-temporal interactions between stimuli can alter the perceived curvature along the outline of a shape (Habak *et al.*, 2004, 2006). To better understand these interactions, we used a forward and backward masking paradigm with Radial Frequency (RF) contours while measuring RF detection thresholds. In Experiment 1, we presented a mask alongside a target contour, and altered the stimulus onset asynchrony (SOA) between this target-mask pair and a temporal mask. We found that a temporal mask increased thresholds when it preceded the target-mask stimulus by 130-180 ms, but decreased thresholds when it followed the target-stimulus mask by 180 ms. Furthermore, Experiment 2 demonstrated that the effects of temporal and spatial masks are approximately additive. We discuss these findings in relation to theories of transient and sustained channels in vision.

5.2   Introduction

It is generally believed that local features encoded in early visual cortical areas (i.e., V1, V2) are integrated in extrastriate areas to form increasingly complex visual representations (Kourtzi *et al.*, 2003; Ostwald *et al.*, 2008; Van Essen *et al.*, 1992; Wilson and Wilkinson, 2015). The last two decades has seen an advancement in understanding how mid-level visual areas combine low-level information to form representations of extended curves and simple shapes, but the majority of this work has used static contours (see Loffler, 2008, 2015 for reviews). Given that neurons throughout the visual pathway integrate information across space and time (Breitmeyer and Ganz, 1977; Hess *et al.*, 2003; Lamme and Roelfsema, 2000; Tanskanen *et al.*, 2008), it is important to understand how mid-level representations may be altered by spatio-temporal interactions arising between shapes.

Radial Frequency (RF) contours have been used by many visual researchers to probe mid-level representations of curvature along closed contours (Bell and Badcock, 2009; Bell *et al.*, 2009; Habak *et al.*, 2004, 2006; Jeffrey *et al.*, 2002; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998). These stimuli are useful because they provide an easy way to manipulate features, such as curvature and angular frequency of curvature extrema, that drive population responses in area V4 of macaques (Gallant *et al.*, 1996; Pasupathy and Connor, 1999, 2001, 2002). Compound forms of these stimuli also can be used to represent outlines of more complex shapes and objects, and thus are useful in deconstructing complex forms into simpler components that are easier to study (Loffler *et al.*, 2005; Wilkinson *et al.*, 1998; Wilson and Wilkinson, 2002; Wilson *et al.*, 2000).

Several studies have demonstrated that spatial interactions between adjacent

shapes can increase RF detection thresholds (Bell *et al.*, 2007; Habak *et al.*, 2004, 2009). In the one study that examined temporal interactions between shapes, Habak *et al.* (2006) found that RF detection thresholds along RF contours were elevated significantly by the onset of a mask presented approximately 80-110 ms after the onset of the target, and that shapes presented after the first backward mask did not increase the magnitude of masking. However, it remains unclear how spatial and temporal masking interact. Given that visual scenes typically consist of multiple moving objects, it is important to understand how spatial positioning of objects might contribute to disruptions in processing of curvature information over time. Therefore, the aim of the current study is to examine how spatial interactions might modulate dynamic processes involved in the perception of closed-contour shapes.

In Experiment 1, we use a spatio-temporal masking paradigm similar to Habak *et al.* (2006) while measuring RF detection thresholds. A spatial mask always appeared alongside targets, while a temporal mask was presented at one of eleven stimulus onset asynchronies (SOAs) relative to target-mask pairs. In Experiment 2, we examined possible spatio-temporal interactions by measuring temporal masking with and without a spatial mask. Results across both experiments show that the spatial mask has an additive masking effect beyond that evoked by a temporal mask alone. However, the additive effect between spatial and temporal masks is observed only when the temporal mask precedes the spatial mask. We discuss these findings in relation to theories of sustained and transient channels of vision.

5.2.1 Methods

Participants

Two of the authors (MS and CQD), and one naïve (VAL), experienced psychophysical observer participated in Experiment 1. The mean age of the observers was 24.3 years (age range: 22-26) and all observers had normal or corrected-to-normal visual acuity. Experimental protocols were approved by the McMaster University Research Ethics Board, and consent of the participant was collected prior to the start of the experiment.

Apparatus and stimuli

Stimuli were generated in Matlab 7.10.0 (Mathworks Inc., Natick, MA) on an iMac 3GHz Quad-core Intel Xeon computer, and were displayed using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The visual display was a NEC Multisync FE700+ graphics monitor with a pixel resolution of 1024×768 (86 pixels/deg) and a refresh rate of $60\,\mathrm{Hz}$. The display had a mean luminance of $35.2\ \mathrm{cd/m^2}$ and was the only light source in the room. Stimuli were viewed binocularly at a distance of $131\,\mathrm{cm}$, which was maintained through the use of a chinrest. From this distance, a single pixel subtended $41.5\,\mathrm{arcsec}$.

Radial frequency contours were generated by sinusoidally modulating the radius of a circle in polar coordinates according to the equation

$$r(\theta) = \bar{r}(1 + A sin(\omega\theta + \phi)) \tag{5.1}$$

where θ is the angle in radians, $\bar{r}$ is the mean radius of the contour, A is the amplitude

of modulation expressed as a proportion of the radius of the circle, ω is the radial frequency in cycles per circumference (cy/2π), and ϕ is angular phase (Wilkinson *et al.*, 1998). The modulation amplitude, *A*, determines the magnitude of curvature: values could range between 0 and 1 to prohibit overlap between the RF contour and the polar center. [1]

The cross-sectional luminance profile was defined by a fourth derivative of a Gaussian (see Wilkinson *et al.*, 1998), with a peak spatial frequency of 8 cpd and a luminance contrast of 99%. Across all experimental conditions, target contours had a radial frequency of 5 and a mean radius of 1.14°. Baseline thresholds were measured in displays that contained only the target contour (i.e., no spatial or temporal mask) at three phases (45°, 135°, and 225°). Modulation amplitudes of RF spatial and temporal masks were set to 15 times the baseline detection thresholds because masks of higher modulation amplitude elicit stronger masking effects relative to masks appearing near threshold (Habak *et al.*, 2004, 2006). The mean target-mask distance was 0.60°.

The phase of spatial and temporal masks was always 0°, whereas the target-mask phase could be 0°, 90°, or 180°. Varying the target-mask phase served two purposes. First, varying phase introduces spatial uncertainty about the location of deformed segments along target shapes. Spatial uncertainty promotes stronger summation of

---

[1]It is important to note that the relation between contour curvature and modulation amplitude is not straightforward. For example, a common definition of curvature is the rate of change of orientation with respect to polar angle (Dickinson *et al.*, 2013; Pasupathy and Connor, 2001, 2002; Poirier and Wilson, 2006), which is a continuous quantity that depends on several parameters in addition to modulation amplitude (e.g., radial frequency). Nevertheless, increasing modulation amplitude results in an average increase in contour curvature when other RF parameters remain fixed. Therefore, for the purposes of reporting changes in curvature sensitivity with a single number, in the remainder of this article we will express curvature in terms of modulation amplitude, with the understanding that this quantity does not represent curvature at specific polar angles along the RF contour.

curvature signals (Green *et al.*, 2018a,b,c), presumably because the uncertainty forces observers to monitor multiple spatial locations along a contour onscreen in order to detect modulations in curvature (Green *et al.*, 2018a). Thus, masking effects that are a consequence of combining curvature signals along a contour should be amplified by varying phase across trials. Second, previous research has shown that masking strength is approximately linearly related to the phase alignment between two shapes (Habak *et al.*, 2004, 2006). Therefore, we can examine the effect of phase on masking strength across SOA conditions, albeit at a group level using amalgamated data (see Result sections for more detail).

To evaluate how dynamic interactions affect spatial interactions between shapes, one spatial mask was always presented alongside a target contour, while a temporal mask was presented at one of eleven stimulus onset asynchronies (SOAs) relative to the target-mask pair (± 280 ms, ± 230 ms, ± 180 ms, ± 130 ms, ± 80 ms, and 0 ms). At an SOA of 0 ms, the contrasts of the spatial and temporal masks were summed and averaged. Because the contrast of each mask was set to the same value of 99%, the effect of performing this averaging procedure is equivalent to using only a single spatial mask appearing alongside the RF target.

Psychophysical Procedure

Detection thresholds were measured using the method of constant stimuli and a 2-IFC paradigm. Prior to the start of the experiment, observers underwent a 60 s light adaptation period, during which the observer fixated the center of the display, followed by practice trials with auditory feedback to ensure they were familiar with the stimuli and task. Observers initiated each trial by pressing the spacebar on a

computer keyboard. Upon initiating a trial, a fixation dot flickered onscreen for 50 ms and was followed, after a delay of 200 ms, by the presentation of two stimulus intervals that were separated by an inter-stimulus interval of 700 ms (Figure 5.1A). In backward masking conditions, each interval began with the presentation of a target stimulus that was followed by the presentation of a temporal mask. In forward masking conditions, each interval began with the presentation of a temporal mask followed by a target stimulus. In one interval the target stimulus was a contour deformed according to Equation 5.1 and in the other interval the stimulus was a comparison contour which was a circle. Observers were asked to identify which of the two intervals contained a deformed target contour by pressing one of two keys on the keyboard. A spatial mask contour appeared concurrently with the target and comparison contours. The duration of each temporal mask and each target/comparison stimulus was 30 ms. The SOA between the temporal mask and target stimulus, which was one of the eleven SOAs listed above, remained the same within a block of trials. Across trials and intervals, stimuli were spatially jittered 0.17° in a random direction from the center of the screen. Figure 5.1 illustrates a typical sequence of events for both forward and backward masking conditions.

A single experimental session contained seven different radial modulation amplitudes that were shown 30 times in random order, for a total of 210 trials per block. The three target-mask relative phase combinations (0°, 90°, and 180°) were randomly interleaved within a given block. Experimental sessions consisted of either 5 or 6 blocks, and a minimum of 8 experimental sessions were completed to ensure two thresholds were recorded at each SOA condition. Each session took approximately 1.5 hours to complete.
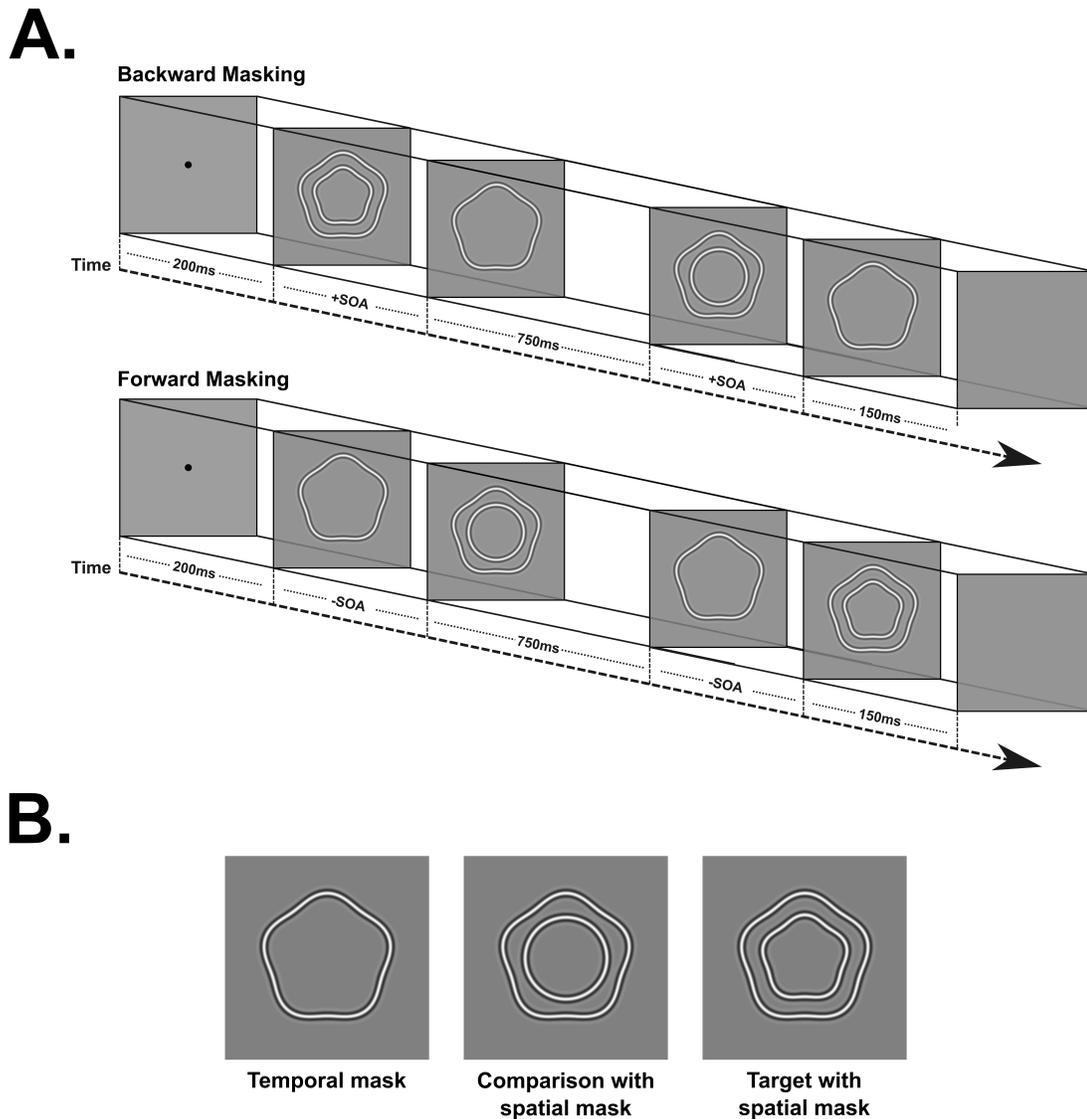
Figure 5.1: A) A typical sequence of events for both forward and backward masking conditions. Temporal offset between the presentation of a mask and a target-mask pair was varied across eleven SOAs (± 280, ± 230, ± 180, ± 130, ± 80, 0 ms). B) Example of RF contours used in Experiment 1. Target and mask contours are shown at 10% modulation amplitude for illustrative purposes.

Data analysis

The statistical computing software R was used to perform all analyses reported within this paper (R Core Team, 2017). Data for each block were fit using maximum likelihood estimation with a psychometric function defined as:

$$\psi(x; \alpha, \beta, \gamma, \lambda) = \gamma + (1 - \gamma - \lambda)F_W(x; \alpha, \beta) \qquad (5.2)$$

where x is RF modulation amplitude, λ is the lapse rate, γ the guess rate, and $F_W(x; \alpha, \beta)$ is a Weibull function (Weibull, 1951) defined as

$$F_W = 1 - \exp(-(\frac{x}{\alpha})^\beta) \qquad (5.3)$$

Threshold was defined as the RF modulation amplitude yielding 75% detection accuracy.

Thresholds were analyzed using a mixed linear model estimated with the lme4 package (Bates *et al.*, 2015). Degrees of freedom for the mixed model were approximated using the Kenward-Roger method (Kenward and Roger, 1997), as this method rescales the F ratios in addition to adjusting the degrees of freedom to better approximate F-distributions for mixed linear models (Judd *et al.*, 2012). For brevity, we report only the F tests from the linear mixed-effects regression analyses (i.e., Analysis of Variance of Type III sums of squares with Kenward-Roger approximation for degrees of freedom). Post-hoc comparisons were performed using paired, two-tailed t tests with p values adjusted with the Holm-Bonferonni method and familywise α equal to 0.05 (Holm, 1979) unless otherwise stated.

5.2.2   Results

Figure 5.2 plots detection thresholds for each observer and group-averaged thresholds as a function of SOA between a temporal mask and target-mask pair. In general, RF detection thresholds were elevated relative to baseline measures across all SOAs tested, and forward masks elicited greater masking than backward masks (see Figure 5.3). For forward masks, a clear peak in masking occurred at SOAs of either 130 or 180 ms in each observer. No consistent peak was found in the backward masking conditions; however, a consistent improvement – a dip in the masking function – was found at 180 ms. Results for Experiment 1 are summarized in Table 5.1, where peak masking in the backward- and forward-masking conditions is listed for each observer.

These observations were confirmed statistically through the use of a mixed linear model. Because the zero SOA condition did not contain a temporal mask, it was omitted from these analyses. The mixed model was fitted with two fixed effects (Mask-Type and SOA), and two random effects (Observer and Session). The analysis revealed a significant interaction between Mask-Type and SOA ($F_{4,59.03} = 8.40$, $p <$ .0001).

To further investigate the cause of the interaction, two separate mixed models, one for each masking condition (Backward and Forward) were performed. Each model was fitted with a fixed effect of SOA, and two random effects (Observer and Session). For backward masks, the main effect of SOA was significant ($F_{4,26.42} = 3.32$, $p = $ .025). Likewise, for forward masks, there was also significant main effect of SOA ($F_{4,26.24} = $ 6.38, $p = 0.001$). The significant main effects of SOA were analyzed with follow-up pairwise comparisons. For backward masking conditions, post-hoc comparisons revealed differences between thresholds obtained at an SOA of 180 ms and thresholds
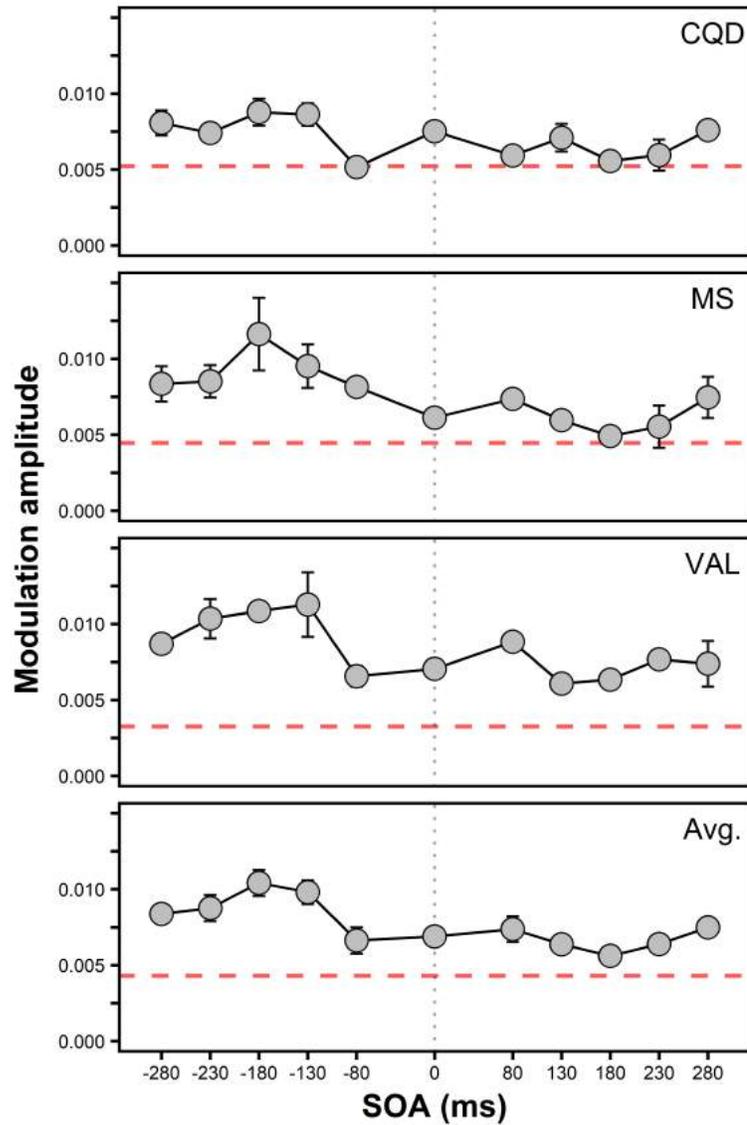
Figure 5.2: Results from Experiment 1. RF detection thresholds are plotted as a function of the SOA between a mask and target-mask pair. The dashed horizontal lines represent baseline thresholds measured in the absence of spatial and temporal masks. Note that only a spatial mask was present in the 0 SOA condition. Errors bars represent $\pm 1\,\mathrm{SEM}$.
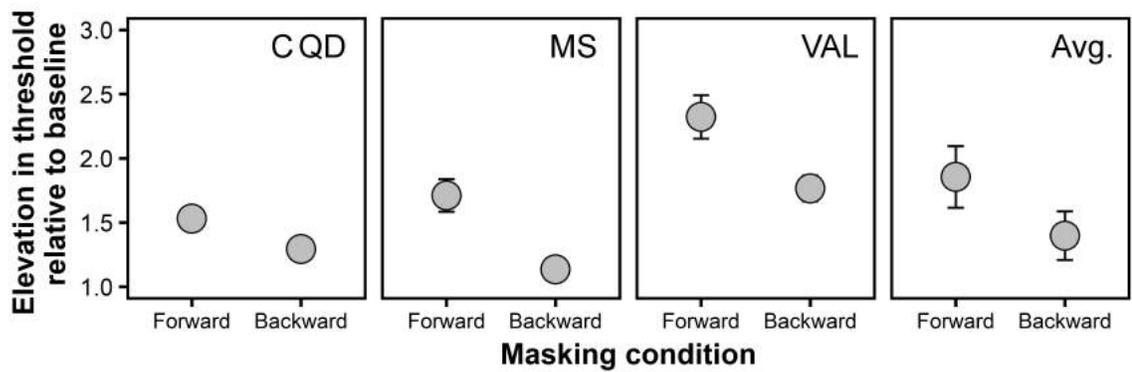
Figure 5.3: Average elevation in detection thresholds relative to baseline for both forward and backward masking conditions, where thresholds were collapsed across SOAs. For all three observers, strength of masking is greater in the forward relative to the backward masking condition, as shown by larger elevations in detection thresholds compared to baseline performance. Error bars represent $\pm 1\,\mathrm{SEM}$.

Table 5.1: SOAs for forward and backward masks that led to peak elevations in detection thresholds.

Observers	Peak Negative SOA	Peak Positive SOA
MS	180 ms	280 ms
CQD	180 ms	280 ms
VAL	130 ms	80 ms
Average	163.33 ms	213.33 ms

at SOAs of 80 ms and 280 ms SOA conditions ($p_{Holm} < .05$). For forward masks, post-hoc comparisons revealed significant differences between thresholds obtained at an SOA of -80 ms and thresholds at SOAs of -130 ms and -180 ms SOA conditions ($p_{Holm} < .01$).

To examine in greater detail how masking strength changes as a function of SOA between forward and backward masks, difference scores were computed at corresponding SOAs between the two conditions, as seen in Figure 5.4. At short and long SOAs, the strength of masking between conditions remained similar, as evidenced by difference scores near zero. However, the magnitude of masking differed at intermediate SOAs, as forward masks exerted greater elevations in threshold compared to backward masks. These data suggest that a temporal mask affects spatial interactions between adjacent shapes primarily at SOAs of $\approx$ 180 ms. It should be noted that masking increased as SOA increased from -80 to -180 ms, but *decreased* as SOA increased from 80 to 180 ms (see Figure 5.2), which serves to further exaggerate the difference between thresholds at SOAs of $\pm$180 ms. However, this result only reinforces our main finding that temporal masks interact with target shapes at both positive and negative SOAs at approximately 180 ms, albeit in potentially different ways. Pairwise t tests were used to determine whether thresholds at SOAs of $\pm$180 ms differ from the zero SOA condition where only a spatial mask is present: A significant difference was found between thresholds for the zero and -180 ms ($t_{(7)} = 3.20$, $p = .015$) and 180 ms ($t_{(7)} = -3.24$, $p = .014$) SOA conditions.

To visualize the effect of phase at each SOA, data for each observer were collapsed across sessions prior to fitting a psychometric function to ensure reliable estimates of threshold were obtained. Data were then averaged across observers at each SOA for
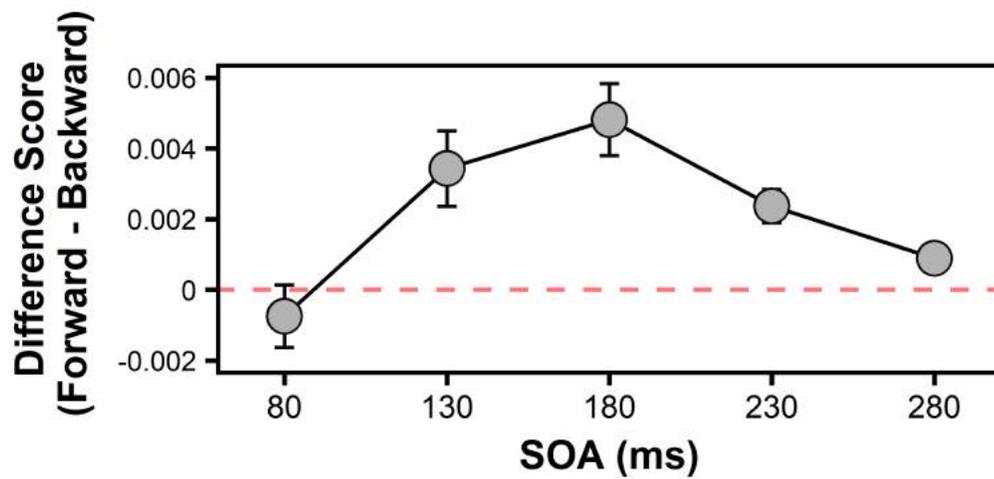
Figure 5.4: The difference between thresholds measured with forward and backward masks is plotted as a function of SOA. The dashed horizontal line represents a difference score of zero: points falling above the line indicate conditions in which thresholds were higher with forward than backward masks. Thresholds measured with forward masks generally were higher than thresholds measured with backward masks, with the largest difference occurring at an SOA of 180 ms. Error bars represent ± 1 SEM.

each target-mask phase combination. As seen in Figure 5.5, at all SOAs the largest elevations in threshold occurred when target and masks were phase aligned (i.e., 0°), and essentially no masking occurred when the target-mask phase was 180°. Also, the difference between forward and backward masking conditions was much greater in phase-aligned conditions.

Experiment 1 demonstrates that a temporal mask can modulate the effect of a spatial mask on RF detection thresholds. Compared to thresholds obtained with only a spatial mask, thresholds decrease when a temporal mask is presented approximately 180 ms after the target-mask pair and increase if the temporal mask precedes the target-mask pair by 180 ms. Consistent with previous masking studies, the largest elevations in threshold are observed when the target and spatial mask are phase aligned (i.e., 0°). To further investigate how the presence of a spatial mask affects temporal masking, in Experiment 2 we measured the magnitude of forward masking with and without spatial masks.

5.3   Experiment 2

5.3.1   Methods

Participants

Six young adults participated in Experiment 2 ($M = 24.50$ years; $SD = 3.45$, range: 20-26). Three observers (CQD, MS, VAL) had participated in Experiment 1, and three (EAM, LUC & CIV) were new, experienced psychophysical observers. All
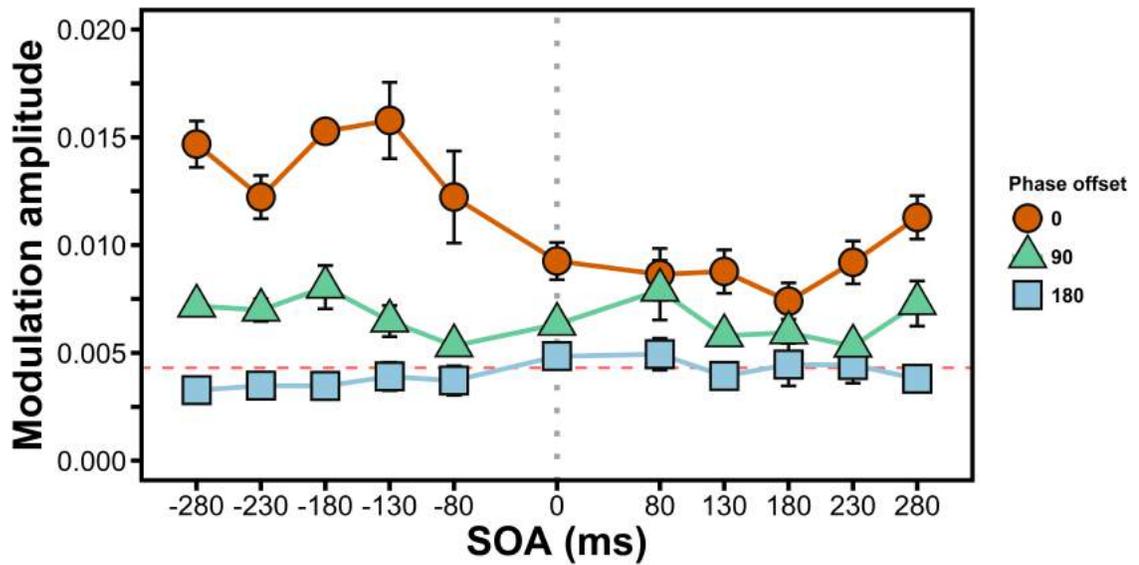
Figure 5.5: Averaged RF detection thresholds with each target-mask phase combination plotted as a function of SOA. The dashed horizontal line represents the baseline threshold (averaged across observers) measured in the absence of spatial and temporal masks. Note that only a spatial mask was present in the 0 SOA condition. Errors bars represent ± 1 SEM.

participants, excluding the main author (MS), were naïve with regards to the experimental hypotheses, and had normal or corrected-to-normal visual acuity. Experimental protocols were approved by the McMaster University Research Ethics Board, and consent of the participant was collected prior to the start of the experiment.

Stimuli, Apparatus, & Procedure

The stimuli, apparatus, and psychophysical procedure were the same as those used in Experiment 1 except for the following differences. First, all observers except CIV were tested with only a subset of forward (and zero) SOA conditions (0, -80 ms, -130 ms, -180 ms) in two masking conditions (i.e., spatial mask present vs. absent), for a total of eight experimental conditions (4 SOAs × 2 mask combinations). In conditions that used only a temporal mask, the zero SOA condition was identical to the baseline condition in which no mask appeared alongside the target contour. The sequence of events on each trial for spatial mask present and absent conditions is illustrated in Figure 5.6.

A single experimental session contained seven different radial modulation amplitudes that were shown 30 times in random order, for a total of 210 trials per block. Experimental sessions consisted of 8 blocks of trials, one block for each experimental condition, with block sequence randomized within a session. A minimum of 2 experimental sessions were completed to ensure at least two thresholds were recorded in each condition. Each session took approximately 2.5 hours to complete.

After completing the main experiment, we measured thresholds in three observers (MS, VAL, and CIV) using SOAs of ±80 ms, where only a temporal mask was present. Each observer ran in a minimum of two sessions per condition, yielding two thresholds
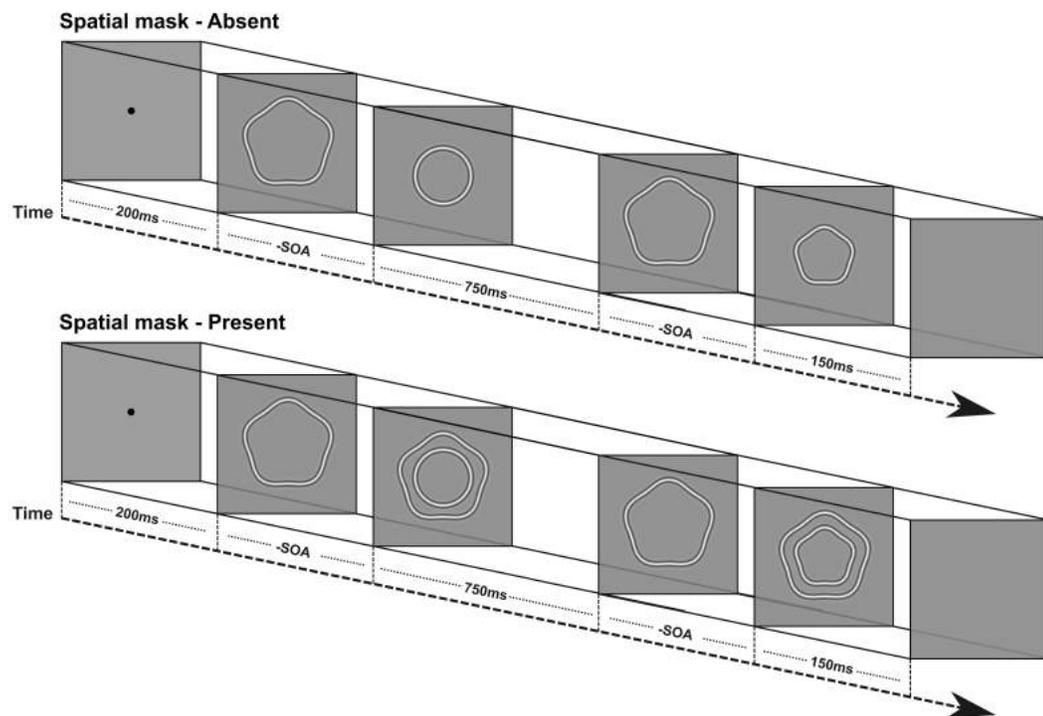
Figure 5.6: An illustration of the sequence of events during a trial in the spatial mask absent (top) and spatial mask present (bottom) conditions in Experiment 2. The SOA between the target and temporal mask was -180, -130, -80, or 0 ms. Target and mask contours are shown at 10% modulation amplitude for illustrative purposes.

per observer per condition. Each session took approximately 1 hour to complete.

5.3.2   Results

Results for Experiment 2 are shown in Figure 5.7. Masking functions varied across observers, especially in conditions in which a spatial mask was present. For most observers, elevations in detection thresholds were largest at the largest SOAs tested, where the onset of the temporal mask far preceded that of target shapes. Furthermore, masking strength was greatest when a spatial mask was present compared to thresholds observed when absent, and the effect of the spatial mask was, on average, similar across SOAs.

The data in Figure 5.7 were analyzed with a linear mixed-effects model, with Spatial Mask (Presence vs. Absence) and SOA as two fixed effects, and two random effects (Observer and Session). The ANOVA revealed a significant main effect of Spatial Mask ($F_{1,52.35} = 23.94$, $p < .0001$) and SOA ($F_{2,49.11} = 9.41$, $p < .001$). The interaction between Spatial Mask Presence and SOA was not significant ($F_{2,49.11} = 1.52$, $p = .23$).

Post-hoc comparisons using the Holm-Bonferonni procedure were performed between SOAs collapsed across masking conditions. Comparisons between thresholds measured with SOAs of 130 and 180 ms and other SOAs tested were significant ($p_{Holm} < .01$ in all cases), but the comparison between 130 and 180 ms were not significant ($p_{Holm} = .31$).

Experiment 1 found that a temporal mask presented 80 ms after a target contour produced minimal masking when a spatial mask was presented simultaneously with the target. This result is surprising because Habak *et al.* (2006) found that a temporal
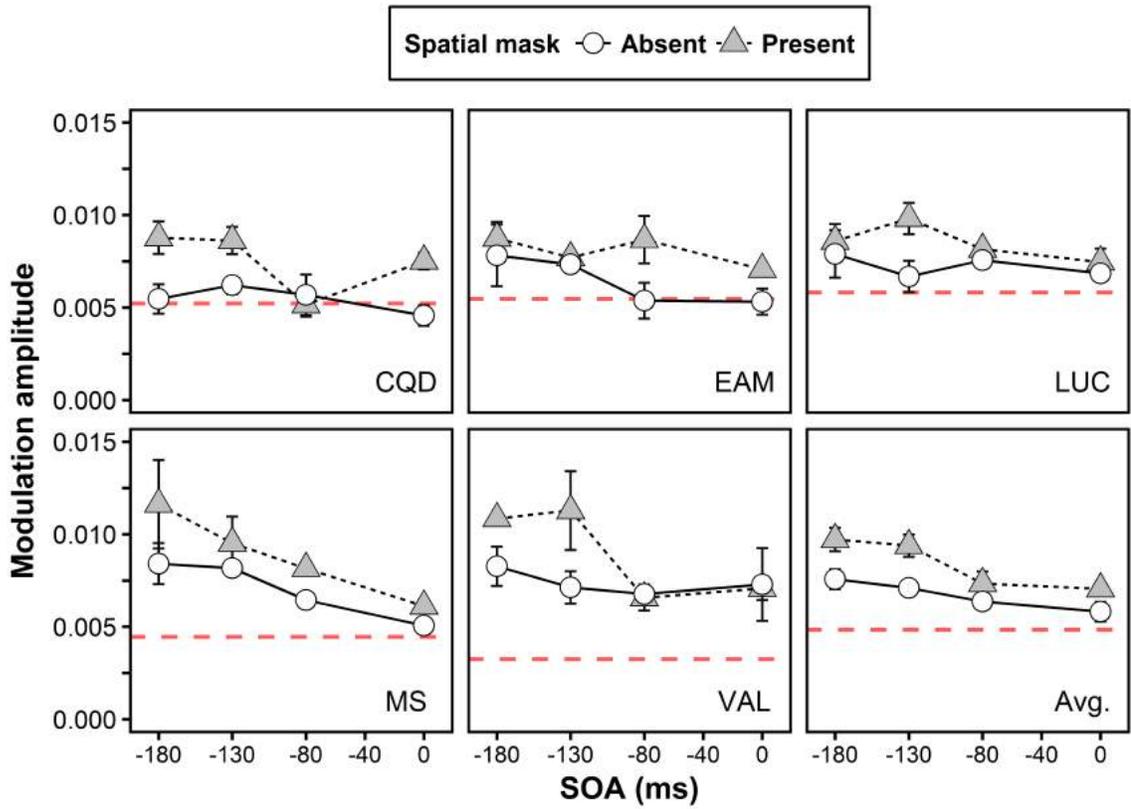
Figure 5.7: Results for Experiment 2 where masking is evaluated at only negative and zero SOAs in the presence (dotted) and absence (solid) of spatial masks. Average RF detection threshold plotted as a function of SOA of a temporally offset mask. The dashed horizontal line represents the baseline detection thresholds, with errors bars representing ±1 SEM.

mask at that SOA produced significant masking when a spatial mask was not present. To test whether removing the spatial mask would increase the effect of backward masks, we measured thresholds in three observers (MS, VAL, CIV) with temporal masks at SOAs of ±80 ms. We selected 80 ms because that SOA was near the peak of the masking function observed by Habak *et al.* (2006) for backward masks when no spatial mask was present. In our experiment, we found that backward masking was so great that some observers still did not reach ceiling performance at the highest amplitude modulations tested. Therefore, we fit psychometric functions to responses averaged across observers. The fitted psychometric functions, which can be seen in Figure 5.8, indicate that backward masking (SOA = 80 ms) was significantly greater than forward masking (SOA = -80 ms). That fact that significant masking occurred at an SOA of 80 ms is consistent with the results of Habak *et al.*.

As in Experiment 1, we examined the effect of phase by estimating thresholds for each target-mask phase combination at each SOA. To increase the reliability of our thresholds estimates, we collapsed data across test sessions before fitting psychometric functions to the data from each observer, and we then averaged thresholds across observers in each condition. Figure 5.9 shows that masking in the spatial-mask present and absent conditions was greatest when the target-mask phase was 0°, and that essentially no masking occurred when the phase was 180°. Also, the effect of SOA was much larger when the phase offset was 0°.

Our results demonstrate that the presence of a spatial mask increases RF detection thresholds relative to thresholds obtained with a temporal mask, and that the effect of the spatial mask did not (on average) differ significantly across SOAs. These results are consistent with the hypothesis that masking between shapes is the sum
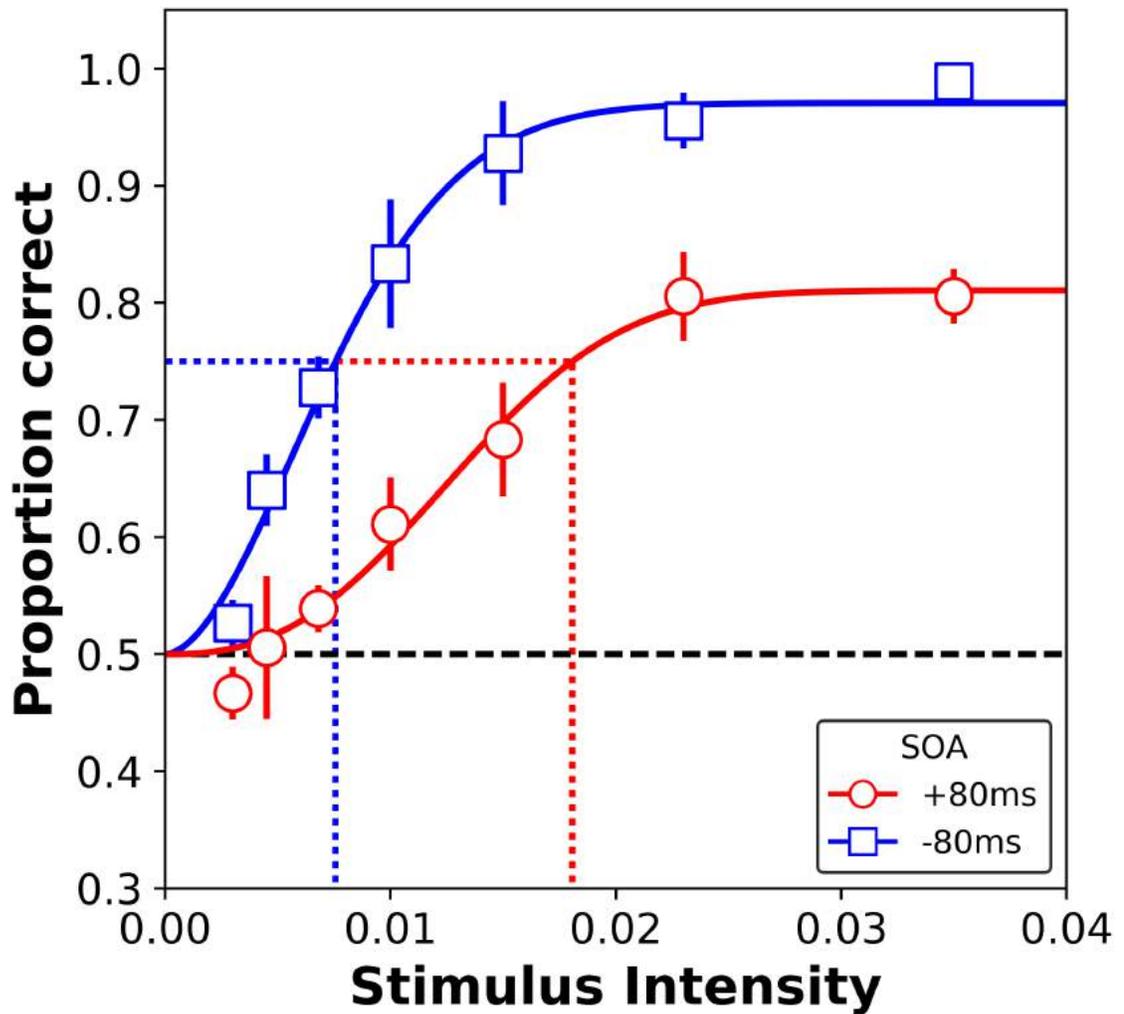
Figure 5.8: Psychometric functions fit to RF detection data averaged across three observers for conditions where temporal masks (with no spatial mask) appeared at SOAs of ± 80 ms. Errors bars represent ± 1 SEM.
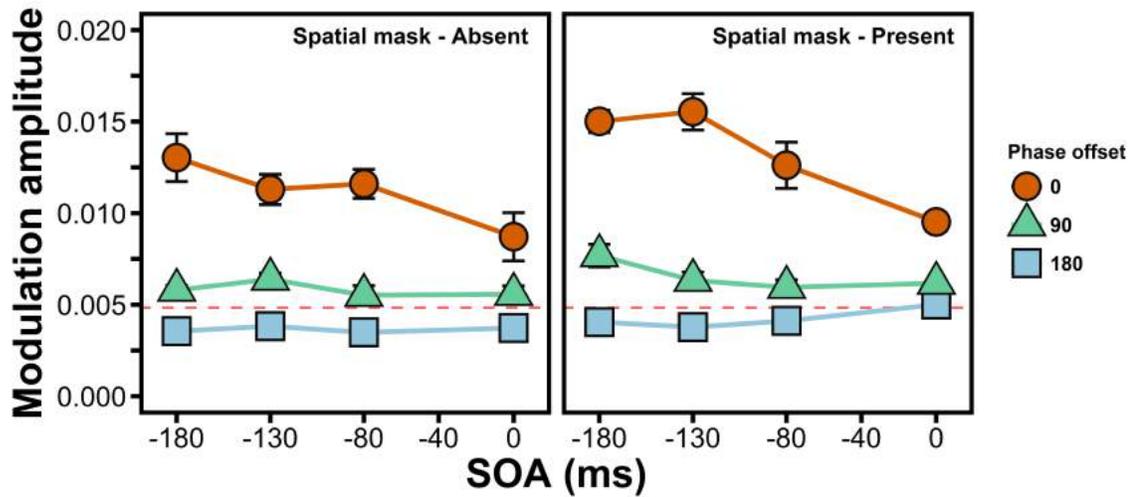
Figure 5.9: Average threshold for each target-mask phase combination plotted as a function of SOA for spatial mask present (right panel) and absent (left panel) conditions in Experiment 2. The dashed horizontal line represents the baseline detection threshold (averaged across observers). Errors bars represent $\pm 1\,\mathrm{SEM}$. In cases where error bars are not visible, the standard error was smaller than the width of the symbols.

of two components: one due to the effect of a temporal mask and another due to the presence of a spatial mask. One important caveat is that our data suggest that potential interactions between spatial and temporal masking may differ significantly across observers (a point that we discuss further in Section 5.4.6). Also, as previously observed, elevations in threshold were largest for phase aligned patterns, with thresholds approaching baseline values for conditions with increasing degrees of target-mask phase offset.

5.4 Discussion

The current study investigated the effect of spatial masks on temporal interactions between shapes. Experiment 1 found that RF detection thresholds measured with forward, temporal masks elevated thresholds relative to thresholds obtained with a spatial mask alone. Conversely, thresholds measured with backward masks improved slightly relative to conditions obtained with a spatial mask alone. Experiment 2 found that masking produced by a mask that precedes the target, and a mask presented concurrently with the target, are (to a first approximation) additive. In other words, forward masking reflected approximately the additive effects of the temporal and spatial masks. Lastly, in both experiments masking strength was strongly dependent on the phase alignment of patterns, with masking decreasing with with increasing phase offset between target and masking shapes. Overall, our results demonstrate that interactions between shapes can be explained by a simple additive model with static and dynamic components that are strongly modulated by phase.

Habak *et al.* (2006) argued that *backward* masking between shapes is driven by the first mask transient: when a stimulus sequence consisted of multiple masks presented after the target, only the first temporal mask produced significant masking. However, Habak *et al.* (2006) only examined the influence of a backward mask on spatial interactions between shapes across a limited range of SOAs (i.e., 80 ms-110 ms). Experiment 1 demonstrated that backward masks presented 180 ms after the target-mask stimulus lower thresholds compared to conditions where the temporal mask is absent. Furthermore, our results from Experiments 1 and 2 suggest that for *forward* masking conditions, the first mask is not solely responsible for driving the magnitude of masking observed. Instead, spatial masks presented concurrently alongside a target contour can further elevate RF detection thresholds beyond the effect evoked by a temporal mask alone. Below, we discuss plausible origins for each shape masking component in relation to theories of sustained and transient channels of vision.

5.4.1 Timescale of information processing along the visual hierarchy

The time needed for visual information to evoke neuronal responses (i.e., mean cortical latencies) differs dramatically across visual cortical areas (Capalbo *et al.*, 2008; Lamme and Roelfsema, 2000; Nowak and Bullier, 1997; Schmolesky *et al.*, 1998). For the purposes of our discussion, we are interested in the mean latency it takes for mid-level visual areas to process and encode representations of curvature along closed contours. An area of the visual cortex that has received much attention for encoding simple representations of shape is V4, as studies in macaques suggest that neurons within V4 produce population responses that represent curvatures at different polar

angles relative to the center of a closed contour (Gallant *et al.*, 1996; Pasupathy and Connor, 1999, 2001, 2002). This population code within V4 can also be used to reconstruct simple shapes (Pasupathy and Connor, 2001, 2002), and therefore can potentially serve as a foundation for building more complex representations of form at higher level cortical areas along the visual processing hierarchy (e.g., IT).

Physiological studies suggest that the mean response latency of neurons within V4 is approximately 100 ms after the onset of a visual stimulus (Capalbo *et al.*, 2008; Lamme and Roelfsema, 2000; Nowak and Bullier, 1997; Schmolesky *et al.*, 1998). Given that the response latency of V4 neurons is approximately 56 ms (Lee *et al.*, 2007), then this latency roughly corresponds to the SOA where peak masking ($\approx$ 130-180 ms) was observed in Experiment 1. This correspondence between maximal masking and V4 mean response latencies is consistent with the idea that V4 is a candidate area for processing shapes, and that masking between shapes may reflect interactions among V4 neurons. To begin our discussion on how temporal and spatial factors might manifest within V4, we must briefly review theories of sustained and transient channels in vision (Kulikowski and Tolhurst, 1973).

5.4.2   Sustained and transient channels

Kulikowski and Tolhurst (1973) first postulated that visual information is processed by sustained and transient visual channels. Sustained channels partake in prolonged processing of visual information, and exhibit a long response latency after brief stimulation (Breitmeyer, 1992; Breitmeyer and Ganz, 1977; Breitmeyer and Ogmen, 2000). These channels are hypothesized to play a critical role in processing object features, as prolonged responses are associated with generation of featural

codes (Breitmeyer, 1992; Breitmeyer and Ganz, 1977; Breitmeyer and Ogmen, 2000). In contrast, transient channels are characterized by brief, rapid activation that reorients attention towards different spatial locations and to moving stimuli (Breitmeyer, 1992; Breitmeyer and Ganz, 1977; Breitmeyer and Ogmen, 2000). Properties of sustained and transient channels are similiar to neurons in, respectively, the parvocellular and magnocellular pathways in old and new world monkeys (Kremers, 1998; Nealey and Maunsell, 1994; Ogmen *et al.*, 2003; Yabuta and Callaway, 1998), and those pathways have been considered by some as the neural correlates of sustained and transient channels in humans (Ogmen *et al.*, 2003).

Although sustained and transient channels serve different functional roles in processing information, reciprocal inhibitory interactions between channels affords opportunities for inter-channel crosstalk to occur (Breitmeyer, 1992; Breitmeyer and Ganz, 1977; Breitmeyer and Ogmen, 2000; Ogmen *et al.*, 2003). Inter-channel inhibition is characterized by a sudden termination or degradation in sustained responses to a stimulus by the activity evoked from processing of a new stimulus onscreen (Breitmeyer, 1992; Breitmeyer and Ganz, 1977; Breitmeyer and Ogmen, 2000; Ogmen *et al.*, 2003). Intra-channel interactions are thought to differ from inter-channel inhibition that arises between stimuli that are close together in time, because behavioural studies have found evidence for intra-channel inhibition even when stimuli are separated by long temporal gaps (Breitmeyer, 1992; Breitmeyer and Ganz, 1977; Breitmeyer and Ogmen, 2000; Ogmen *et al.*, 2003). Although theories of intra- and inter-channel inhibition between sustained and transient channels are based mostly upon behavioural evidence, such inhibitory interactions do have plausible neurophysiological correlates (Kruse and Eckhorn, 1996; Yabuta and Callaway, 1998).

Disruption in sustained patterns of activation via intra-channel or inter-channel inhibition produce characteristically different masking functions (Breitmeyer, 1992; Breitmeyer and Ganz, 1977; Breitmeyer and Ogmen, 2000; Ogmen *et al.*, 2003). Type A masking functions, which are thought to reflect intra-channel inhibition, are obtained when temporal offsets between the presentation of a target and mask are small (Breitmeyer and Ganz, 1977; Ogmen *et al.*, 2003). Type A masking functions describe patterns of masking whereby thresholds are most elevated at SOAs occurring proximal to the onset of a target stimulus. At longer intervals between target and mask onset, transient responses from the second stimulus are able to interfere with sustained responses evoked by the first stimulus at higher levels along the visual hierarchy (Ogmen *et al.*, 2003). This interference results in inter-channel inhibition and a so-called Type B masking function, in which masking occurs only at longer temporal offsets between two stimuli (Breitmeyer, 1992; Breitmeyer and Ganz, 1977; Breitmeyer and Ogmen, 2000; Ogmen *et al.*, 2003).

Type A functions do a better job of describing the interference observed in Experiment 1 with backward masks, whereas Type B functions are better for describing the results obtained with forward mask. One possible reason why Type B masking functions are observed for forward masking conditions is that transient responses evoked by a target-mask pair disrupts sustained processing of forward masks within mid-level visual areas (i.e., V4, LOC). A similar pattern of masking was observed by Habak *et al.* (2006), who showed that detection of curvature along a target shape was severely impaired when a mask was presented approximately 80-110 ms after onset of the target. Instead of a mask transient degrading the discriminability of a target, our experiments suggest that the transient response evoked by a target-mask pair

interacts with sustained processing of a temporal mask (see Figure 5.10). Our results from Experiments 1 and 2 are consistent with the idea that interruptions in sustained processing produce masking at longer SOAs.

5.4.3   Spatial masks modulate inter-channel inhibition

In our current study, we assume that temporal masks affect perception of target shapes in one of two ways: 1) Evoked transient responses interfere with sustained processing of similar information upstream; and 2) Sustained processing of a mask is disrupted by downstream transients evoked from stimuli appearing later in sequence. Both of these are examples of inter-channel inhibition. Although explanations based on ideas of inter-channel inhibition may explain why peak forward masking is observed at longer SOAs, it does not explain why the effect of a spatial mask on target discriminability is approximately constant across SOAs tested, as seen in Experiment 2. Transient responses evoked by temporal masks likely do not activate regions along the visual pathway at the same time as target contours, unless there is enough time between stimulus presentations for the transient to affect a region where sustained processing of similar visual information is observed. Therefore, we suggest that intra-channel inhibition is needed to account for the fact that spatial masks further elevating detection thresholds beyond that attributable to the presence of a temporal mask.

Spatial masks interfere with target encoding at different levels of the visual hierarchy over the same time course at which target shapes are processed. In primary visual cortex, interference between a spatial mask and target shape are hypothesized to arise via weak local inhibition between orientation selective filters (Poirier and
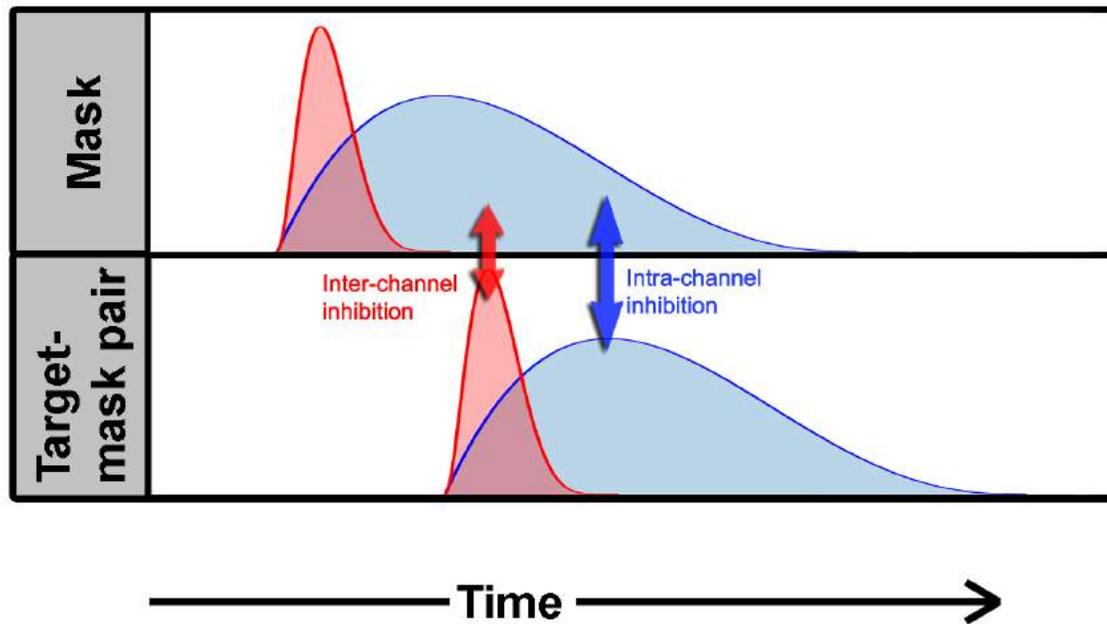
Figure 5.10: Adapted from Ogmen *et al.* (2003). The height of each curve represents the strength of activation evoked from either a mask or target-mask pair in transient (red) or sustained (blue) channels. Transient responses interfere with sustained processing of shape information via inter-channel inhibition, resulting in the termination of processing of information along sustained channels. Intra-channel inhibition also occurs between sustained channels, but this type of interference is negligible in our Experiments given the larger delays between target-mask pairs and temporal masks.

Wilson, 2006, 2007). Inhibition between local filters that process contour orientation results in reduced neural responses, thus propagating weaker signals to areas further upstream that integrate such information to represent shape (Poirier and Wilson, 2007). At higher visual areas, such as V4, interference between spatially adjacent shapes is attributed to the improper summation of curvature signals along a target and mask contours at locations of peak curvature (Habak *et al.*, 2004; Poirier and Wilson, 2006). Such theories are based on the response properties of V4 neurons that respond to curvature extrema relative to the center of a visual stimulus (Pasupathy and Connor, 2001, 2002). Although such encoding schemes confer benefits in representing shapes, such as scale-invariance, if curvature extrema occur at the same polar angle relative to the center of a shape, the response of neurons may saturate, resulting in an impaired ability to encode changes in curvature. Consistent with this idea, Experiments 1 and 2 found that masking is greater in conditions where curvature extrema are aligned compared to conditions where curvature extrema are misaligned, a result that is consistent with the results obtained in previous visual masking studies (Habak *et al.*, 2004, 2006). Taken together, studies of spatial masking between shapes suggest that interference likely arises from interactions arising within the same visual area, and possibly along a similar time-course.

Results from Experiment 2 are consistent with the idea that spatial masks are interfering with processing within the same visual areas in which target shapes are encoded. However, it should be emphasized that intra-channel inhibition arising from spatial interactions between shapes is likely not restricted to a specific cortical area or level of processing, but occurs at different tiers along the visual hierarchy. Furthermore, spatial masking, which we have argued arises from intra-channel inhibition,

appears to have an additive effect on temporal masking. As demonstrated in Experiment 2, a spatial mask only scaled thresholds by a constant factor across all SOAs tested, and therefore did not interact with components of masking that are attributed to temporal masks. This result suggests that the mechanism(s) responsible for dynamic interactions between shapes, such as inter-channel inhibition, is (are) likely separate from those mechanisms, such as intra-channel inhibition, that contribute to static spatial interactions between shapes. Thus, although both intra- and inter-channel inhibition likely affect target visibility when both a spatial and temporal mask are present, their effects are additive rather than multiplicative.

5.4.4   Alternatives to curvature-based interpretations

Although our discussion has focused on the role of curvature in discriminating shapes, our data are agnostic regarding the specific feature being used to discriminate two RF contours. Indeed, recent research has questioned whether changes in the rate of curvature modulation is the critical feature used by observers to discriminate different shapes (Dickinson *et al.*, 2012, 2013, 2015, 2018; Schmidtmann and Kingdom, 2017). For example, Schmidtmann and Kingdom (2017) described a two-stage model in which the difference between maximum and minimum curvature is the primary computation used to code for shape, not curvature change *per se*. Such theories of shape encoding based on the periodicity of curvature extrema are consistent with the results from Experiments 1 and 2, as offsets in phase between two RF contours resulted in less masking.

Based on neuroimaging studies that suggest radial and concentric gratings evoke strongest responses in intermediate visual areas (Gallant *et al.*, 1996; Wilkinson *et al.*,

189

2000), we speculate many shape-specific interactions originate in area V4. Nevertheless, it is plausible that disruptions in shape discrimination originate in other visual areas. Such a claim is supported by recent work by Salmela *et al.* (2016) that investigated shape representations using fMRI and multi-voxel pattern analysis. Despite the advantage conferred in using multi-voxel analysis in decoding multidimensional patterns of activation relative to older methods (Norman *et al.*, 2006), Salmela *et al.* (2016) failed to find pattern selectivity for radial frequencies in area V4, although RF specific patterns of activation were found in other intermediate cortical areas (e.g., V3d, IPS0). However, Salmela *et al.* (2016) did observe that RF contours elicited strong activation within area V4, and suggested that the neuroimaging device used (i.e., fMRI) may lack the resolution needed to discriminate between neural patterns generated by different RF contours (Salmela *et al.*, 2016). Thus, area V4 still remains an important candidate area to consider in contributing to the ability to distinguish between shapes.

In identifying visual areas that contribute to accurate discrimination of shapes, consideration of the task used to probe these areas is paramount. Although the spatial summation of curvature signals has been well studied (Green *et al.*, 2017, 2018a,b,c; Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012), far less attention has been devoted to understanding how curvature signals are processed over prolonged durations (Green *et al.*, 2018b; Habak *et al.*, 2006). This is problematic because much of our understanding of shape perception may be limited to feedforward processes that occur following a brief stimulus presentation and fail to consider how shape perception is affected by recurrent cortical processes connections to shape processing that are known to that operate over longer durations (Lamme and

Roelfsema, 2000). Therefore, additional research is needed to better understand the time course of recruitment of visual areas in passing diagnostic information regarding shape identity.

5.4.5   Constraints on models of shape perception

Our data suggest that the mechanisms responsible for discriminating shapes operate over a prolonged time-course. In considering a purely feedforward model of visual processing along the ventral stream, our results are consistent with the hypothesis shape discrimination depends on computations performed in intermediate visual areas such as V4, which tend to have moderate to long response latencies to visual stimuli (Lamme and Roelfsema, 2000). Our results also are consistent with results reported by Habak *et al.* (2006) who demonstrated that backward masking was greatest at longer SOAs (approximately 80-110 ms).

Together, these results place important constraints on the time course over which shape processing unfolds. For example, our results pose a problem for theories that propose shape discrimination depends on analysis only of low-level orientation filters (Baldwin *et al.*, 2016), as activity in primary visual cortex typically decays rapidly after initial stimulus onset (Celebrini *et al.*, 1993). Evidence from single-cell recordings of neurons in macaque inferior temporal cortex also suggest that the spike train of these neurons fail to encode information specific to shape identity until after approximately 20 ms (Kovács *et al.*, 1995). If shape discrimination depended only on analysis of output from V1 orientation selective neurons, the largest interactions between shapes should occur at short SOAs, a prediction that is inconsistent with our data. However, our results do not rule out the possibility that extrastriate areas are

performing local computations over large spatial regions (Baldwin *et al.*, 2016).

5.4.6   Individual variability in the pattern of masking

For Experiments 1 and 2, there was notable inter-observer variability in the pattern of masking observed. Given the multitude of factors that influence masking, it is not surprising that these functions can differ between observers. Masking a stimulus introduces noise to the signal (i.e., target) either at early levels of visual processing, such as those arising from spatial masks, or along higher levels of visual processing typically evoked by temporal masks (Enns and Di Lollo, 2000). The addition of this sensory noise can make it difficult for a decision making process to categorize a stimulus based on this noisey visual representation, and consequently this uncertainty is reflected as additional noise in the behavioural responses of an observer. Furthermore, the level of attention directed by an observer to a stimulus can also impact levels of performance. The role of covert attention has been shown to affect the perception of a target in a variety of visual tasks (see Posner and Petersen, 1990, for a review). Therefore, additions of noise at sensory encoding, along with how these noisey representations affect higher cognitive processes (e.g., decision and/or attention) at least partly explain why masking functions are so variable across observers. However, more research is needed in understanding how uncertainty via noise, sensory or otherwise, can affect human perception of shape.

5.5   Conclusion

To summarize, the aim of the current study was to investigate the impact of spatial over dynamic interactions between shapes in multi-shape displays. Results from Experiment 1 demonstrate that backward and forward masks have very different effects on RF detection. Specifically, the effect of a forward mask depended significantly on SOA, with peak masking occurring for SOAs near -180 ms, whereas as backward masks had a smaller effect , with thresholds improving at a SOA of 180 ms. Results from Experiment 2 demonstrate that the presence of a spatial mask serves to increase thresholds above those obtained with a backward mask alone, and that the effect of the spatial mask was nearly constant across SOA. Across both experiments, masking strength was strongly modulated by phase alignment between target and masking patterns. Overall, these results suggest that spatial and temporal masks contribute to masking in RF detection tasks, and that their effects are (to a first approximation) additive. Our results are consistent with theories of inhibition between transient-sustained channels of vision, where both intra- and inter-channel inhibition are considered in describing spatio-temporal interactions between multiple shapes. Future work should aim to further elucidate whether interactions between shapes attributed to static and dynamic components of shape processing are additive, or whether the mechanisms governing shape interactions may operate under non-additive routines.

5.6   References

Baldwin, A. S., Schmidtmann, G., Kingdom, F. A. A., and Hess, R. F. (2016). Rejecting probability summation for radial frequency patterns, not so Quick! *Vision Research*, **122**, 124–134.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48.

Bell, J. and Badcock, D. R. (2009). Narrow-band radial frequency shape channels revealed by sub-threshold summation. *Vision Research*, **49**(8), 843–50.

Bell, J., Badcock, D. R., Wilson, H., and Wilkinson, F. (2007). Detection of shape in radial frequency contours: independence of local and global form information. *Vision Research*, **47**(11), 1518–22.

Bell, J., Wilkinson, F., Wilson, H. R., Loffler, G., and Badcock, D. R. (2009). Radial frequency adaptation reveals interacting contour shape channels. *Vision Research*, **49**(18), 2306–17.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433–36.

Breitmeyer, B. G. (1992). Parallel Processing in Human Vision: History, Review, and Critique. *Advances in Psychology*, **86**, 37–78.

Breitmeyer, B. G. and Ganz, L. (1977). Temporal studies with flashed gratings: Inferences about human transient and sustained channels. *Vision Research*, **17**(7), 861 – 865.

Breitmeyer, B. G. and Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, **62**(8), 1572–1595.

Capalbo, M., Postma, E., and Goebel, R. (2008). Combining Structural Connectivity and Response Latencies to Model the Structure of the Visual System. *PLOS Computational Biology*, **4**(8), e1000159.

Celebrini, S., Thorpe, S., Trotter, Y., and Imbert, M. (1993). Dynamics of orientation coding in area V1 of the awake primate. *Visual Neuroscience*, **10**(5), 811–825.

Dickinson, J. E., Harman, C., Tan, O., Almeida, R. A., and Badcock, D. R. (2012). Local contextual interactions can result in global shape misperception. *Journal of Vision*, **12**(11), 3.

Dickinson, J. E., Bell, J., and Badcock, D. R. (2013). Near their thresholds for detection, shapes are discriminated by the angular separation of their corners. *PLoS ONE*, **8**(5), 1–9.

Dickinson, J. E., Cribb, S. J., Riddell, H., and Badcock, D. R. (2015). Tolerance for local and global differences in the integration of shape information. *Journal of Vision*, **15**(3), 21–21.

Dickinson, J. E., Haley, K., Bowden, V. K., and Badcock, D. R. (2018). Visual search reveals a critical component to shape. *Journal of Vision*, **18**(2), 2–2.

Enns, J. T. and Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, **4**(9), 345–352.

Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W., and Van Essen, D. C. (1996). Neural responses to polar, hyperbolic, and cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*, **76**(4), 2718–2739.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2017). Global processing of random-phase radial frequency patterns but not modulated lines. *Journal of Vision*, **17**(9), 18–18.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018a). Convergent evidence for global processing of shape. *Journal of Vision*, **18**(7), 7–7.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018b). The effect of spatiotemporal displacement on the integration of shape information. *Journal of Vision*, **18**(5), 4–4.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018c). Integration of shape information occurs around closed contours but not across them. *Journal of Vision*, **18**(5), 6–6.

Habak, C., Wilkinson, F., Zakher, B., and Wilson, H. R. (2004). Curvature population coding for complex shapes in human vision. *Vision Research*, **44**(24), 2815–23.

Habak, C., Wilkinson, F., and Wilson, H. R. (2006). Dynamics of shape interaction in human vision. *Vision Research*, **46**(26), 4305–20.

Habak, C., Wilkinson, F., and Wilson, H. R. (2009). Preservation of shape discrimination in aging. *Journal of Vision*, **9**(12), 18.1–8.

Hess, R. F., Wang, Y. Z., and Dakin, S. C. (1999). Are judgements of circularity local or global? *Vision Research*, **39**(26), 4354–60.

Hess, R. F., Hayes, A., and Field, D. J. (2003). Contour integration and cortical processing. *Journal of Physiology-Paris*, **97**(2–3), 105–119.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**(3), 65–70.

Jeffrey, B. G., Wang, Y. Z., and Birch, E. E. (2002). Circular contour frequency in shape discrimination. *Vision Research*, **42**(25), 2773–9.

Judd, C. M., Westfall, J., and Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, **103**(1), 54–69.

Kenward, M. G. and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, **53**(3), 983–997.

Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., and Logothetis, N. K. (2003). Integration of local features into global shapes: Monkey and human fmri studies. *Neuron*, **37**(2), 333–346.

Kovács, G., Vogels, R., and Orban, G. A. (1995). Cortical correlate of pattern backward masking. *Proceedings of the National Academy of Sciences*, **92**(12), 5587–5591.

Kremers, J. (1998). Spatial and temporal response properties of the major retino-geniculate pathways of Old and New World monkeys. *Documenta Ophthalmologica*, **95**(3-4), 229–245.

Kruse, W. and Eckhorn, R. (1996). Inhibition of sustained gamma oscillations (35-80 Hz) by fast transient responses in cat visual cortex. *Proceedings of the National Academy of Sciences*, **93**(12), 6112–6117.

Kulikowski, J. J. and Tolhurst, D. J. (1973). Psychophysical evidence for sustained and transient detectors in human vision. *The Journal of Physiology*, **232**(1), 149–162.

Lamme, V. A. F. and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, **23**(11), 571–579.

Lee, J., Williford, T., and Maunsell, J. H. R. (2007). Spatial Attention and the Latency of Neuronal Responses in Macaque Area V4. *Journal of Neuroscience*, **27**(36), 9632–9637.

Loffler, G. (2008). Perception of contours and shapes: Low and intermediate stage mechanisms. *Vision Research*, **48**(20), 2106–2127.

Loffler, G. (2015). Probing intermediate stages of shape processing. *Journal of Vision*, **15**(7), 1–1.

Loffler, G., Wilson, H. R., and Wilkinson, F. (2003). Local and global contributions to shape discrimination. *Vision Research*, **43**(5), 519–30.

Loffler, G., Yourganov, G., Wilkinson, F., and Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience*, **8**(10), 1386–1391.

Nealey, T. A. and Maunsell, J. H. (1994). Magnocellular and parvocellular contributions to the responses of neurons in macaque striate cortex. *Journal of Neuroscience*, **14**(4), 2069–2079.

Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, **10**(9), 424–430.

Nowak, L. G. and Bullier, J. (1997). The Timing of Information Transfer in the Visual System. In K. S. Rockland, J. H. Kaas, and A. Peters, editors, *Extrastriate Cortex in Primates*, number 12 in Cerebral Cortex, pages 205–241. Springer US.

Ogmen, H., Breitmeyer, B. G., and Melvin, R. (2003). The what and where in visual masking. *Vision Research*, **43**(12), 1337–1350.

Ostwald, D., Lam, J. M., Li, S., and Kourtzi, Z. (2008). Neural Coding of Global Form in the Human Visual Cortex. *Journal of Neurophysiology*, **99**(5), 2456–2469.

Pasupathy, A. and Connor, C. E. (1999). Responses to contour features in macaque area v4. *Journal of Neurophysiology*, **82**(5), 2490–2502.

Pasupathy, A. and Connor, C. E. (2001). Shape representation in area v4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, **86**(5), 2505–2519.

Pasupathy, A. and Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience*, **5**(12), 1332—1338.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysocs: transforming numbers into movies. *Spatial Vision*, **10**, 437–42.

Poirier, F. J. and Wilson, H. R. (2006). A biologically plausible model of human radial frequency perception. *Vision Research*, **46**(15), 2443–2455.

Poirier, F. J. and Wilson, H. R. (2007). Object perception and masking: Contributions of sides and convexities. *Vision Research*, **47**(23), 3001–3011.

Posner, M. I. and Petersen, S. E. (1990). The Attention System of the Human Brain. *Annual Review of Neuroscience*, **13**(1), 25–42.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Salmela, V. R., Henriksson, L., and Vanni, S. (2016). Radial Frequency Analysis of Contour Shapes in the Visual Cortex. *PLOS Computational Biology*, **12**(2), e1004719.

Schmidtmann, G. and Kingdom, F. A. A. (2017). Nothing more than a pair of curvatures: A common mechanism for the detection of both radial and non-radial frequency patterns. *Vision Research*, **134**, 18–25.

Schmidtmann, G., Kennedy, G. J., Orbach, H. S., and Loffler, G. (2012). Non-linear global pooling in the discrimination of circular and non-circular shapes. *Vision Research*, **62**, 44–56.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., and Leventhal, A. G. (1998). Signal Timing Across the Macaque Visual System. *Journal of Neurophysiology*, **79**(6), 3272–3278.

Tanskanen, T., Saarinen, J., Parkkonen, L., and Hari, R. (2008). From local to global: Cortical dynamics of contour integration. *Journal of Vision*, **8**(7), 15–15.

Van Essen, D. C., Anderson, C. H., and Felleman, D. J. (1992). Information Processing in the Primate Visual System: An Integrated Systems Perspective. *Science*, **255**(5043), 419–423.

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, **18**, 292–97.

Wilkinson, F., Wilson, H. R., and Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, **38**(22), 3555–68.

Wilkinson, F., James, T. W., Wilson, H. R., Gati, J. S., Menon, R. S., and Goodale, M. A. (2000). An fMRI study of the selective activation of human extrastriate form vision areas by radial and concentric gratings. *Current Biology*, **10**(22), 1455–1458.

Wilson, H. R. and Wilkinson, F. (2002). Symmetry perception: a novel approach for biological shapes. *Vision Research*, **42**(5), 589–97.

Wilson, H. R. and Wilkinson, F. (2015). From orientations to objects: Configural processing in the ventral stream. *Journal of Vision*, **15**(7), 4.

Wilson, H. R., Wilkinson, F., Lin, L. M., and Castillo, M. (2000). Perception of head orientation. *Vision Research*, **40**(5), 459–72.

Yabuta, N. H. and Callaway, E. M. (1998). Functional Streams and Local Connections of Layer 4c Neurons in Primary Visual Cortex of the Macaque Monkey. *Journal of Neuroscience*, **18**(22), 9489–9499.

Chapter 6

General Discussion

The ability to accurately perceive the shape of object outlines serves as a basis for object recognition. This dissertation predominantly examines how perception of contour curvature along outline shapes changes in multi-shape displays. Previous psychophysical studies have demonstrated the remarkable sensitivity with which the human observers discriminate curvature along the contour of closed contour shapes when a single shape appears onscreen (Green *et al.*, 2017, 2018a; Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann and Kingdom, 2017; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998). However, the perception of curvature in a target contour is disrupted when a second shape appears in close spatial or temporal proximity to the target shape (Habak *et al.*, 2004, 2006, 2009; Poirier and Wilson, 2007). To better understand the factors that modulate interactions between shapes that affect curvature discriminability, the research presented within this dissertation made extensive use of lateral masking paradigms. Across experiments, the spatio-temporal profile of a non-target shape was systemically altered, and its effect on target discriminability measured. This chapter discusses these experimental findings within the context of

our current state of knowledge on shape perception in humans, the limitations of using visual masking as a tool to study shape perception, and questions that remain unanswered about shape perception that future directions of research should aim to explore.

6.1 Properties of shape that modulate interactions across contours

6.1.1 Alignment of curvature extrema

Shape perception is commonly modelled as a process whereby curvature encoding neurons, such as those in V4, respond to concentric forms by pooling changes in orientation across their receptive fields (Habak *et al.*, 2004; Poirier and Wilson, 2006). Physiological data on response properties of V4 neurons provide support for such an encoding scheme of shape, as V4 neurons respond to curvature extrema relative to the center of an object (Carlson *et al.*, 2011; Pasupathy and Connor, 2001, 2002). Although physiological evidence from macaques has provided evidence in support of the use of curvature extrema in coding for shape, visual masking has been a useful tool used by vision researchers to better understand what curvatures features are used by the human visual system to represent shape. The basic idea is that masking between adjacent contours occurs when key features used by the visual system are close together in space and/or time: the fact that the features are in close proximity means that they will likely fall within the same receptive field and therefore will saturate the response of neurons sensitive to that feature. The result is typically an inability to accurately encode those features at the neuronal level, and is reflected

in behavioural measures as elevations in detection or discrimination thresholds (i.e. masking). By varying parameters of the mask that are thought to be most relevant in coding for a stimulus of interest, one can identify the alterations leading to the greatest changes in masking. One can then infer that features that modulate strength of masking are likely important to the visual system in coding for that stimulus.

Although the radial organization of curvature-sensitive cells in V4 allows for a scale invariant representation of shape, this representation scheme should also result in an impaired ability to encode curvatures aligned along the same polar angle, because alignment increases the likelihood that the responses of curvature sensitive neurons saturate. Visual masking studies have provided support for this argument, as masking strength increases with the number of curvature extrema that are radially aligned with a target (Habak *et al.*, 2004), suggesting that the response of V4 neurons may increase as more curvatures fall within their receptive fields. This suggests that masking between contours may result as a consequence of the inappropriate pooling of curvature signals across the receptive field of V4 neurons. If the response function of V4 neurons is modelled as a compressive non-linearity, then subsequent curvatures falling along receptive fields of V4 neurons will need to be of higher amplitude in order to elicit an equivalent change in activation strength (Habak *et al.*, 2004). However, as the distance between positive curvature extrema increases, either by increasing the distance between shapes, or by presenting shapes at relative phase offsets, fewer curvature features should fall within overlapping receptive fields, and therefore sensitivity to curvature should be restored. Habak *et al.* (2004) demonstrated that manipulations that primarily reduced the alignment of positive curvature extrema between target and masking shapes restored sensitivity to curvature along the target shape.

Therefore, results from shape masking studies such as these argue that the human visual system makes use of positive curvature extrema in coding for shape.

Our results from Chapter 2 and 3 are consistent with data from previous shape masking studies that suggest positive curvature extrema are important features used by the visual system to code for shape. As evidenced from our results in Experiment 2 of Chapter 2, the addition of high frequency perturbations along the mask contour, which were otherwise identical in shape to target contours, did not reduce lateral masking effects unless the high frequencies perturbation was significantly above detection threshold. The addition of high frequency curvature along the contour of a mask has two effects: it disrupts orientation and position cues along the shape, and it increases the distance of positive curvature extrema between shapes. Both orientation and position cues (Day and Loffler, 2009), and the distance between curvature extrema (Habak *et al.*, 2004), are thought to be important features used by the visual system in shape encoding. While orientation and position cues may be important overall in coding for shape, our results suggest that disruptions to parallel alignment of orientation and position cues between shapes is not primarily what drives masking between shapes. Instead, it is likely the increase in distance between positive curvature features by increasing the amplitude of the high frequency component that led to less masking. Therefore, our results from Chapter 2 further support the argument that interactions between shapes arises from the inappropriate pooling of curvatures along positive extrema.

Our results from Chapter 3 are also consistent with theories of shape encoding that rely upon representing curvatures at different polar angles. In Experiments 1 and 3 of Chapter 3, when two otherwise identical shapes are aligned in phase, discrimination

thresholds for curvature along the target shape are elevated compared to baseline measures collected with no mask present. However, as the phase offset between the target and mask increases, the strength of masking declines, and typically thresholds are restored to levels of sensitivity observed for baseline conditions. Phase offsetting shapes ensures curvature extrema do not fall within receptive fields of curvature sensitive cells responding to the same curvature polarity. However, models of shape perception typically only attribute masking to arise from the parallel alignment of positive curvature maxima, not minima, between two shapes (Habak *et al.*, 2004; Poirier and Wilson, 2006). Therefore, to test whether alignment of curvature extrema of both polarities results in masking, we used half-wave rectified masks of both polarity in Experiment 3 of Chapter 3. Our results demonstrate that half-wave rectified masks of both polarity elevate detection threshold when aligned in phase with target shapes, with masking strength declining as the phase offset between target and mask decline. These data argue that the alignment of both curvature extrema can modulate the response of V4 neurons, and therefore models of shape perception should incorporate curvatures of both polarity in encoding schemes for shape.

However, consistent with current models of shape perception, our results from Chapter 4 demonstrate that observers are uniformly sensitive to curvature deformations as a function of polar angle. Although not explicitly investigated in multi-shape displays, these results suggest that interactions arising from the alignment of curvature extrema will be of similar magnitude despite the absolute angular position at which curvature deformations occur. In other words, so long as laterally displaced curvature extrema are in alignment, the effect of masking should be the same despite a rotation of both shapes, assuming both shapes are rotated in the same direction (i.e.,

clockwise or counter-clockwise), and of equal magnitude. Furthermore, these data argue that relative angular offsets between closed contours with similar curvature profiles are what allows the visual system to discriminate between two spatially adjacent shapes, as neural mechanisms sensitive to curvature deformations along closed contours are agnostic with regards to absolute angular position.

6.1.2 Radial frequency and phase

In addition to the effects of local curvature structure, masking has been shown to be influenced by more complex shape features that can only be computed by integrating visual information across extended spatial scales (Habak *et al.*, 2004; Poirier and Wilson, 2007). One such feature, radial frequency (cyl/2π), has received special attention because psychophysical studies suggest that curvature detection of radial frequencies between 2 and 8 cycles/2π is based on the integration of curvature along extended parts of an enclosed contour (Green *et al.*, 2017, 2018a; Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998). Specifically, for 2-8 cycles/2π improve with increasing number of modulation cycles greater than is predicted by probability summation (Hess *et al.*, 1999; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012). Although the relation between curvature detection thresholds and radial frequency has been thoroughly examined in displays containing a single shape (Bell and Badcock, 2009; Bell *et al.*, 2008, 2009; Hess *et al.*, 1999; Kurki *et al.*, 2014; Loffler *et al.*, 2003; Schmidtmann *et al.*, 2012; Wilkinson *et al.*, 1998), little is known about how differences in the ability of shape mechanisms to integrate curvature frequency might affect other important features that modulate contextual interactions between shapes.

Chapter 3 begins to address this gap in the literature by examining how radial frequency affects spatial interactions based on the phase alignment between two shapes. Habak *et al.* (2004) demonstrated that masking between shapes of the same radial frequency is greatest when shapes are perfectly aligned (0° phase offset)and declines monotonically with increasing phase offset. The effect of phase alignment is consistent with the predictions of computational models of the response properties of V4 neurons, which attribute between-shape masking to the rapid saturation in response of V4 neurons to positive curvatures falling within their respective receptive fields (Habak *et al.*, 2004, 2009; Poirier and Wilson, 2006). Because radial frequency affects the ability of shape mechanisms to integrate local curvatures along the contour of shapes, and phase-dependent interactions are thought to arise from the improper summation of curvature signals, a plausible prediction is that phase-dependent interactions differ as a function of curvature frequency. However, the experiments in Chapter 3 demonstrated that the effect of phase alignment on between-shape masking did not depend significantly on the radial frequency of the mask.

The results from Chapter 3 are consistent with the idea that phase-dependent masking occurs *after* contour curvature is encoded by multiple, radial-frequency selective channels (Bell and Badcock, 2009; Bell *et al.*, 2007, 2009; Habak *et al.*, 2004). As previously modelled by Poirier and Wilson (2006), the visual system may be performing a Fourier-like decomposition of curvature along the outline contour of shapes. Evidence from psychophysical studies suggest such a computational approach may be made possible via the existence of shape channels that encode curvature frequencies between approximately 3 - 10 cycles/2π (Bell and Badcock, 2009; Bell *et al.*, 2007, 2009; Habak *et al.*, 2004). Each shape channel codes for the power at which a given

curvature frequency occurs along the contour of a shape (Bell and Badcock, 2009; Bell *et al.*, 2007, 2009; Habak *et al.*, 2004). The existence of shape channels would allow the visual system to compute the similarity between two shapes, such as a target and mask, using something akin to computing the dot product between vectors encoding for each shape. Such an encoding scheme would be useful in differentiating between shapes to devote additional resources in representing each shape independently. Based on this argument, phase selective masking between two contours results from a computation judging the similarity between two shapes. If shapes are similar enough, the visual system may try to compress representations by collapsing across lower visual areas that pool for curvature. Alternatively, as the similarity between shapes decreases with increasing phase offset, the visual system begins treating curvatures occurring along the contour of each shape as belonging to separate objects, and masking declines. Such computations of shape similarity would also result in the effect of phase remaining the same across radial frequencies. Our results from Chapter 1 are consistent with this framework. In those experiments, we found that adding low- and medium-amplitude high frequency curvatures along to an otherwise low frequency mask did not affect between-shape masking. The failure to find an effect of the high-frequency component is difficult to explain if contour curvature is processed by a single, broad-band channel. At very high modulation amplitudes, the added high frequency component did affect masking, presumably by altering other important properties governing interactions between shapes (i.e., distance between curvature extrema).

Another possibility is that our results from Chapter 3 are not the result of interactions between frequency-selective shape channels, nor differences in summation

strength of curvature signals, but instead arise as a consequence of the coding scheme used by curvature-sensitive neurons in visual cortex. Models of shape perception typically are based on the population responses of V4 neurons that represent the polar angles of curvature extrema along a closed-contour shape (Carlson *et al.*, 2011; Pasupathy and Connor, 2001, 2002). If curvature is defined as the rate of change in orientation of a tangent line relative to circularity, a simple bi-dimensional population code can represent shapes as magnitudes of deformation from circularity as a function polar angle (relative to the centre of the shape). Adopting such a model of curvature encoding along closed contours leads to the interesting prediction that phase-dependent interactions between shapes depends upon the resolution at which curvature sensitive neurons represent phase shifts across radial frequencies.

In this framework, curvatures occurring at low radial frequencies would elicit a population response from V4 neurons that has a broad spatial profile, in which the neural response would change slowly as a function of polar angle. Conversely, shapes comprised of high curvature frequencies would elicit responses with a narrow spatial profile, in which the neural response would change rapidly as a function of polar angle. With shifts in phase, the population response will be very similar for those patterns that evoke a spatially spread response, as the level of activation across the population surface will remain similar with small shifts in phase. Therefore, higher level decision masking processes that might be monitoring the output of such population responses would have trouble using the difference map in population responses to judge low radial frequencies were phase shifted. If population responses of V4 neurons are less informative to observers in making discriminatory judgments regarding shifts in phase of shapes, perhaps phase-dependent interactions depend on observer'

sensitivity to shifts in phase of shapes. Data from Chapter 3 are consistent with this hypothesis, as observers were less sensitive to shifts in phase with increasing radial frequency. However, using a computational model of population responses of neurons in V4, we demonstrated that sensitivity to phase shifts does not seem to be well captured by the discriminability of difference maps in the population response elicited by shapes of different radial frequency. Difference maps of higher radial frequency contained strong differences in the pattern of peak activation that should be helpful in determining whether these shapes were shifted in phase. Therefore, our results argue that the population response from curvature sensitive neurons likely does not limit phase discrimination performance across radial frequencies. In summary, phase selective masking does not depend on radial frequency, but radial frequency modulates observers' levels of sensitivity to phase shifts of shape, as seen in Chapter 3.

While data from Chapter 3 suggest that population responses from curvature selective units do not limit phase discrimination performance, this result holds only under the assumption that these units are uniformly sensitive to curvature deformations as a function of polar angle, regardless of shape frequency. Otherwise, it would be difficult to predict whether a simple shift in phase might elicit a degenerate activation map that limits performance based on the failure to code for curvatures as a function of polar angle, rather producing activation maps too similar to discriminate (i.e., non-informative difference maps). To test this assumption, Experiments 1 and 2 of Chapter 4 measured curvature detection thresholds across 24 polar angles for two closed contour shapes with curvature deformations of differing angular frequency. Results from Chapter 4 demonstrate that curvature thresholds remain largely invariant

to polar angle, but are linearly scaled based on angular frequency. Thus, phase discrimination thresholds are not limited by the capacity of the visual system in coding for curvature as a function of polar angle, but rather, are limited as a consequence of other factors based on the underlying scheme used by neural units to code for shape.

In considering the importance of curvature polarity in coding for shape, evidence from behavioural studies suggests that positive curvatures are more important in shape encoding than are negative curvatures (Dickinson *et al.*, 2013, 2015, 2018; Loffler *et al.*, 2003). Following this logic, the response properties of V4 neurons have often been modelled as responding exclusively to positive curvature deformations (Habak *et al.*, 2004; Poirier and Wilson, 2006). Physiological investigations into the response properties of curvature sensitive neurons within V4 demonstrate these neurons are respond to both positive and negative curvatures (Pasupathy and Connor, 1999, 2001, 2002). To test the relative contributions of both curvature polarities to patterns of phase-dependent masking, in Experiment 3 of Chapter 3 we used positive and negative half-wave rectified RF contours as lateral masks while measuring thresholds for sensitivity to curvature along a target contour. Consistent with physiological models of V4 neurons, our results argue for the joint contribution of curvature polarities in modulating the level of interaction between spatially adjacent shapes. Although masks comprised of only positive curvature extrema caused greater elevations in thresholds relative to masks sampling only negative curvature extrema, both masking conditions caused smaller elevations in threshold compared to conditions in which both curvature extrema were visible.

One possibility as to why positive curvature extrema, relative to negative extrema, cause greater elevations in threshold is because there exists a bias to use positive over

negative curved extrema to detect deformations along a concentric stimulus (Loffler *et al.*, 2003). Lateral masking of positive curvature maxima would therefore result in larger declines in performance relative to masking curvature minima, as these features are more task relevant to observers in making accurate shape discrimination judgements. These predictions are consistent with our results from Chapter 3, and suggest that although positive curvatures are more important to observers in discriminating between two shapes, interactions between two spatially adjacent shapes can arise from the alignment of curvature maxima of either polarity.

6.1.3   Spatio-temporal interactions between shapes in vision

Changes in visual input across space and time allows for dynamic interactions to emerge within and between visual areas along the processing hierarchy. A dual-channel theory of visual perception, where visual input is processed along sustained and transient channels, has been widely applied in understanding the types of interactions that can emerge between stimulus features across spatial and temporal scales. Sustained channels often are described as being involved in the slow processing of object features, whereas transient channels are implicated in coarse texture processing and rapid encoding of stimulus location (Breitmeyer and Ogmen, 2000).

Differences in speed of processing along each channel type result in interactions between and within channels arising from activity evoked at different stages of visual processing. Interactions between sustained channels, such as those posited to occur when spatial masks are positioned alongside target shapes, are commonly associated with improper pooling of stimulus features occurring within the same cortical areas.

Conversely, temporal interactions between stimuli are thought to arise from disruptions in sustained processing of visual information from transiently evoked responses occurring later in sequence (Breitmeyer and Ogmen, 2000). Therefore, interactions arising between sustained and transient channels may not originate from activity evoked at the same level of processing.

To better understand the role of transient interactions between shapes, Habak *et al.* (2006) presented observers with sequences of shapes, and assessed levels of interaction through elevations in detection thresholds for curvature along target contours. They found that curvature detection thresholds were most impaired in conditions in which a secondary shape was presented approximately 80-110 ms following target onset. Furthermore, Habak *et al.* demonstrated that it was the onset of the secondary shape, also known as its transient response, that was responsible for producing masking.

Although effects of spatial and temporal interactions have been studied independently using single shape displays (Habak *et al.*, 2004, 2006), less is known about the effect of spatial interactions in modulating temporal processes in coding for shape in multi-shape displays, and served as the motivation for the experiments described in Chapter 5. We used the same masking paradigm as that used by (Habak *et al.*, 2006) except that a spatial mask always appeared alongside targets. Experiment 1 demonstrated that a temporal mask altered thresholds only when it preceded target-mask pairs by 130-180 ms. Furthermore, Experiment 2 showed that the masking effects attributable to the spatial and temporal mask stimuli were approximately additive.

Our results from Chapter 5 are consistent with theories of sustained and transient channels of vision, and suggest interactions attributable to each channel are additive.

Specifically, spatial interactions between adjacent shapes can be attributed to interactions that occur within sustained channels (within the same visual area) that process the mask and target contours. Conversely, temporal interactions between shapes are attributed to interactions within transient channels, where onset of a shape elicits activity in transient channels that disrupt sustained processing of shape information that might have occurred prior in sequence (Habak *et al.*, 2006). Surprisingly, the co-occurence of both types of masking, spatial and dynamic, has a simple linear relation, rather than a more complicated mapping (i.e., exponential). One explanation is that such additive effects are only evident when dynamic interactions between shapes precede spatial interactions between shapes, as masking effects attributable to the spatial mask will only serve to further modulate sustained processing of the temporal mask. In other words, it is only the sustained processing of the temporal mask that is interfering with the sustained processing of the target shape and spatial mask. Although this explanation is consistent with our results obtained in Experiment 1, it fails to consider that a transient response from the onset of the target also interacts with the sustained processing of masks presented earlier in time. Perhaps, this interaction between transient and sustained processes degrades signal fidelity of target shapes, therefore only scaling masking effects caused by spatial interactions between shapes by a constant.

These arguments are consistent with the results obtained in Chapter 5, and also suggest that mechanism(s) responsible for dynamic interactions between shapes likely operate separately from those that contribute to static interactions between shapes. Thus, in the context of shape perception, effects caused by interactions between sustained and transient channels of vision appear to be described by additive rather

than multiplicative operations.

6.1.4 Limitations of visual masking in understanding shape perception

Despite the widespread use of visual masking techniques to study vision, there are inherent limitations in using visual masking to probe mechanisms that code for shape. The first problem arises from the design of these studies, which purposefully elicit unfavourable perceptual outcomes by attempting to degrade target visibility using a visual mask. By varying task difficulty, the visual system has been shown to selectively shift attention towards task-relevant stimulus features (Navalpakkam and Itti, 2005). Therefore, results from visual masking studies that suggest a particular feature may be important for coding a stimulus may be true only for that particular task and experimental design. To further complicate matters, directing attention towards specific stimulus features has a modulatory effect over signal strength in coding for those features within a display (Carrasco *et al.*, 2004; Maunsell and Treue, 2006; Saenz *et al.*, 2002). This might further worsen the problem of stimulus features appearing more important in coding for a stimulus than they might otherwise appear when only a single stimulus is visible onscreen.

Another problem arising from the use of masking paradigms in studying shapes is that thresholds reflect changes in the visibility of a target stimulus, but do not reveal what aspects of a target representation are being degraded. For example, elevations in curvature detection thresholds along a target contour are observed when a shape with similar shape characteristics, such as curvature frequency, are positioned alongside the target. In scenarios like these, elevations in threshold are often attributed to

the improper integration of stimulus features. Such explanations are appealing, but say very little about what is changing about the appearance of a target that results in changes to performance. Perhaps visual filters are pooling curvatures across extended spatial regions, and therefore the final percept is that of a target containing extraneous curvatures beyond those contained in the original input. Alternatively, the visual system may encode a target stimulus with sufficient resolution such that curvatures are represented accurately, but the addition of a nearby stimulus similar in shape to the target might impact how higher-level decision processes interpret population responses that code for shape. Both explanations would result in elevations in threshold being observed, but make different predictions regarding whether the coding scheme that gives rise to the appearance of a target shape is actually altered.

Interpretation of the results of visual masking studies also depends upon whether spatial or temporal masks are used. Different mechanisms are postulated to affect performance when a temporal interval is inserted between the presentation of a mask and target stimulus. When two shapes are presented simultaneously onscreen, the story is more straightforward, as inputs will be processed along stages of the visual pathway along a parallel time course. Thus, visual structures of stimuli can be systematically altered, and effects of these alterations on task performance can be measured. Over the course of many experiments, researchers can start to gain an understanding of the contribution of task relevant visual structures, and how these structures might contribute to the overall percept of a stimulus. Because different visual structures might also selectively activate different visual areas, an understanding of the contribution of different cortical areas in processing a stimulus might also be gained.

However, as the temporal interval separating the onset of a target and mask increases, it remains unclear as to where along the visual pathway interactions between target and mask stimuli might occur. For example, residual feedforward activation evoked from the first stimulus might have any interaction with feedforward activation from a subsequently presented stimulus. Physiological studies examining the temporal decay in activation of visual processing areas demonstrate this to be especially true of early visual areas, where evoked reposes to a visual stimulus begin to decay after only 90ms (Schmolesky *et al.*, 1998; Self *et al.*, 2013). This is still considering interactions within a network that arise only from feedforward connections, and differential response activation profiles. When recurrent connections and feedback are also considered, as have been evidenced from neurophysiological studies in primate visual cortex (Lamme and Roelfsema, 2000), interactions arising between stimuli can involve a combination of factors occurring across different levels of processing, time courses, and between channels of vision.

6.1.5 Future directions of research

Shape processing in multi-shape displays

Shape discrimination has been thoroughly studied in displays containing single shapes (Bell and Badcock, 2009; Bell *et al.*, 2009; Day and Loffler, 2009; Hess *et al.*, 1999; Jeffrey *et al.*, 2002; Kurki *et al.*, 2014; Loffler *et al.*, 2003; Wilkinson *et al.*, 1998), while few studies have examined how curvature perception along shape contours changes in multi-shape displays (Bell *et al.*, 2007; Green *et al.*, 2018b; Habak *et al.*, 2004, 2009; Poirier and Wilson, 2007). Of those studies investigating curvature perception using multiple shapes (Habak *et al.*, 2004, 2006, 2009; Hess *et al.*, 1999;

Poirier and Wilson, 2007), the vast majority make use of visual masking paradigms, where the presence of a spatially adjacent shape typically results in elevations in curvature detection thresholds. As described above, there are many limitations in using visual masking paradigms to better understand visual processing of a stimulus. Therefore, contour perception in multi-shape displays should also be investigated using alternative experimental designs other than visual masking.

For example, using visual search in displays composed of multiple RF contours, Almeida *et al.* (2010a,b) demonstrated that individuals with autistic traits exhibit stronger summation of curvatures across shape contours relative to normal healthy controls. The same visual search task was then used by Dickinson *et al.* (2018) to determine what features are most useful in coding for shape. Results demonstrate that visual search performance was most impaired when shapes shared the same polar angle between curvature maximum. From these results, Dickinson *et al.* (2018) argued that polar angle between curvature extrema is an important shape feature used by the visual system to discriminate between multiple shapes in a display, rather than curvature frequency or polarity. These results highlight the differential role shape features may have in displays that contain more than a single object, and the need to explore what features critically drive performance under different experimental designs.

Models of shape perception and psychophysical measures

Population codes are exploited by the visual system to generate multidimensional shape spaces, where each dimension represents a geometric property of importance (e.g., curvature) (Carlson *et al.*, 2011; Pasupathy and Connor, 2001, 2002). Although

computational models of shape perception have been designed to represent shape using population coding schemes (Kempgens *et al.*, 2013; Poirier and Wilson, 2006), these models tend to deviate from models that are more closely based on the neurophysiological properties of V4 cells in monkeys (Carlson *et al.*, 2011; Pasupathy and Connor, 2001, 2002). These different modelling strategies make the problem of linking physiological states with perceptual outcomes very difficult. Researchers often posit that visual capacities as measured through behavioural performance are limited by neural substrates of the visual system (Teller, 1984). Yet, discussions surrounding the shortcomings in using explicit models of population responses of V4 neurons in accounting for psychophysical data is lacking in the literature.

Do population responses in V4 carry enough information to predict shape discrimination performance? In Chapter 3, we applied a closely approximated model of V4 developed by Pasupathy and Connor (2001, 2002) to see whether sensitivity to shifts in rotation of shape for single-cycle contours could be explained by differences between surface activation patterns generated from such a physiologically inspired model. Our results demonstrate that this simple model based entirely on tuning properties of V4 neurons cannot account for results in our behavioural task, and therefore argues that alternative explanations are needed. Comparisons between such neural models and behavioural data should be made more frequent, as greater insight can be gained as to the computational complexity that is needed to describe perceptual outcomes under different stimulus conditions.

To further expand upon such propositional links, future work should aim to approximate the efficiency at which the visual system is able to combine noisy sources of contour curvature to code for shape. Bayesian models (i.e., ideal observers) serve

as one useful tool to provide quantitative benchmarks of efficiency, as these models optimally combine sources of noisy information to produce the best possible estimate of shape (Deneve and Pouget, 2004; Knill and Pouget, 2004; Körding *et al.*, 2007). Use of such Bayesian models would allow researchers to compare human performance on a shape discrimination task to an ideal observer whose performance is always optimal given the information made available to perform a task (Geisler, 2011). To date, no such studies exist that explicitly compare Bayesian models to human performance on tasks related to the detection of discrimination of shapes.

6.2   Concluding remarks

Shape perception is critical in guiding purposeful interactions with our external environment, as object identity it intrinsically linked with its form. Human visual cortex transforms early visual inputs into more concentrically defined structures of perception, namely curvature. While curvature information is well represented by the visual system along isolated closed contours, contextual interactions between shapes can affect sensitivity of mechanisms of shape at coding curvature information. A central question of this dissertation is what factors modulate the level of interaction between shape representation in visual cortex in multi-shape displays. Research in this dissertation revealed that local orientation and position, curvature frequency, phase-alignment, and temporal sequencing are all factors that can modulate interactions across shapes. This research compliments the burgeoning literature on shape perception and understanding how intermediate levels of visual cortex represent complex forms. Future research should seek to determine the efficiency at which the mammalian visual system is able to represent shape from physiologically plausible

models of intermediate levels of visual processing.

6.3   References

Almeida, R. A., Dickinson, J. E., Maybery, M. T., Badcock, J. C., and Badcock, D. R. (2010a). A new step towards understanding Embedded Figures Test performance in the autism spectrum: The radial frequency search task. *Neuropsychologia*, **48**(2), 374–381.

Almeida, R. A., Dickinson, J. E., Maybery, M. T., Badcock, J. C., and Badcock, D. R. (2010b). Visual search performance in the autism spectrum II: The radial frequency search task with additional segmentation cues. *Neuropsychologia*, **48**(14), 4117–4124.

Bell, J. and Badcock, D. R. (2009). Narrow-band radial frequency shape channels revealed by sub-threshold summation. *Vision Research*, **49**(8), 843–50.

Bell, J., Badcock, D. R., Wilson, H., and Wilkinson, F. (2007). Detection of shape in radial frequency contours: independence of local and global form information. *Vision Research*, **47**(11), 1518–22.

Bell, J., Dickinson, J. E., and Badcock, D. R. (2008). Radial frequency adaptation suggests polar-based coding of local shape cues. *Vision Research*, **48**(21), 2293–2301.

Bell, J., Wilkinson, F., Wilson, H. R., Loffler, G., and Badcock, D. R. (2009). Radial frequency adaptation reveals interacting contour shape channels. *Vision Research*, **49**(18), 2306–17.

Breitmeyer, B. G. and Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, **62**(8), 1572–1595.

Carlson, E. T., Rasquinha, R. J., Zhang, K., and Connor, C. E. (2011). A Sparse Object Coding Scheme in Area V4. *Current Biology*, **21**(4), 288–293.

Carrasco, M., Ling, S., and Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, **7**(3).

Day, M. and Loffler, G. (2009). The role of orientation and position in shape perception. *Journal of Vision*, **9**(10), 14.1–17.

Deneve, S. and Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. *Journal of Physiology-Paris*, **98**(1), 249–258.

Dickinson, J. E., Bell, J., and Badcock, D. R. (2013). Near their thresholds for detection, shapes are discriminated by the angular separation of their corners. *PLoS ONE*, **8**(5), 1–9.

Dickinson, J. E., Cribb, S. J., Riddell, H., and Badcock, D. R. (2015). Tolerance for local and global differences in the integration of shape information. *Journal of Vision*, **15**(3), 21–21.

Dickinson, J. E., Haley, K., Bowden, V. K., and Badcock, D. R. (2018). Visual search reveals a critical component to shape. *Journal of Vision*, **18**(2), 2–2.

Geisler, W. S. (2011). Contributions of ideal observer theory to vision research. *Vision Research*, **51**(7), 771–781.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2017). Global processing of random-phase radial frequency patterns but not modulated lines. *Journal of Vision*, **17**(9), 18–18.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018a). Convergent evidence for global processing of shape. *Journal of Vision*, **18**(7), 7–7.

Green, R. J., Dickinson, J. E., and Badcock, D. R. (2018b). Integration of shape information occurs around closed contours but not across them. *Journal of Vision*, **18**(5), 6–6.

Habak, C., Wilkinson, F., Zakher, B., and Wilson, H. R. (2004). Curvature population coding for complex shapes in human vision. *Vision Research*, **44**(24), 2815–23.

Habak, C., Wilkinson, F., and Wilson, H. R. (2006). Dynamics of shape interaction in human vision. *Vision Research*, **46**(26), 4305–20.

Habak, C., Wilkinson, F., and Wilson, H. R. (2009). Preservation of shape discrimination in aging. *Journal of Vision*, **9**(12), 18.1–8.

Hess, R. F., Wang, Y. Z., and Dakin, S. C. (1999). Are judgements of circularity local or global? *Vision Research*, **39**(26), 4354–60.

Jeffrey, B. G., Wang, Y. Z., and Birch, E. E. (2002). Circular contour frequency in shape discrimination. *Vision Research*, **42**(25), 2773–9.

Kempgens, C., Loffler, G., and Orbach, H. S. (2013). Set-size effects for sampled shapes: experiments and model. *Frontiers in Computational Neuroscience*, **7**.

Knill, D. C. and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, **27**(12), 712–719.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., and Shams, L. (2007). Causal Inference in Multisensory Perception. *PLOS ONE*, **2**(9), e943.

Kurki, I., Saarinen, J., and Hyvarinen, A. (2014). Investigating shape perception by classification images. *Journal of Vision*, **14**(12), 1–19.

Lamme, V. A. F. and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, **23**(11), 571–579.

Loffler, G., Wilson, H. R., and Wilkinson, F. (2003). Local and global contributions to shape discrimination. *Vision Research*, **43**(5), 519–530.

Maunsell, J. H. R. and Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, **29**(6), 317–322.

Navalpakkam, V. and Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, **45**(2), 205–231.

Pasupathy, A. and Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, **82**(5), 2490–2502.

Pasupathy, A. and Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, **86**(5), 2505–2519.

Pasupathy, A. and Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience*, **5**(12), 1332–1338.

Poirier, F. J. and Wilson, H. R. (2006). A biologically plausible model of human radial frequency perception. *Vision Research*, **46**(15), 2443 – 2455.

Poirier, F. J. and Wilson, H. R. (2007). Object perception and masking: Contributions of sides and convexities. *Vision Research*, **47**(23), 3001–3011.

Saenz, M., Buracas, G. T., and Boynton, G. M. (2002). Global Effects of Feature-Based Attention in Human Visual Cortex. *Nature Neuroscience*, **5**(7), 631–632.

Schmidtmann, G. and Kingdom, F. A. A. (2017). Nothing more than a pair of curvatures: A common mechanism for the detection of both radial and non-radial frequency patterns. *Vision Research*, **134**, 18–25.

Schmidtmann, G., Kennedy, G. J., Orbach, H. S., and Loffler, G. (2012). Non-linear global pooling in the discrimination of circular and non-circular shapes. *Vision Research*, **62**, 44–56.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., and Leventhal, A. G. (1998). Signal Timing Across the Macaque Visual System. *Journal of Neurophysiology*, **79**(6), 3272–3278.

Self, M. W., van Kerkoerle, T., Supèr, H., and Roelfsema, P. R. (2013). Distinct Roles of the Cortical Layers of Area V1 in Figure-Ground Segregation. *Current Biology*, **23**(21), 2121–2129.

Teller, D. Y. (1984). Linking propositions. *Vision Research*, **24**(10), 1233–1246.

Wilkinson, F., Wilson, H. R., and Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, **38**(22), 3555–68.