UNIVERSITÀ
DEGLI STUDI
DI PADOVA

University of Padova
Department of General Psychology

Doctoral School in Psychological Science
XXVI Cycle

Perceptual binding of static and dynamic signals: a psychophysical and electrophysiological study on contour integration.

Direttore della Scuola: Ch.ma Prof. Francesca Peressotti

Supervisore: Ch.ma Prof. Clara Casco

Ph.D Candidate:
Rosilari Bellacosa Marotti

Abstract (English)

The present work investigates the mechanisms underpinning the integration of local signals (either local orientations, positions or directions) into whole configurations. The investigation is composed of three studies that try to disentangle the issue using a contour integration paradigm. Each of them focuses on a specific aspect of the problem.

Study one compares two integration models: the first is the well known "association field model", based on local lateral interactions between adjacent receptive fields tuned to similar orientation (in primary visual area V1). The second is a second-stage filter that follows rectification of first-order filters. Study two tests, instead, the idea that a local cooperative system is responsible for the integration of directional signals. In addition it investigates whether such a mechanism could explain the "motion facilitation effect" usually found with dynamic (compared to static contours). Finally, study three extends findings from study two recording Visual Evoked Potentials (VEPs) elicited by static and dynamic contours.

My findings provide support to the idea that a mechanism based on local lateral interactions in V1 could account for the integration of static contours, whereas a local cooperative mechanism could account for the integration of static contours. Moreover, these two mechanisms interact, in a way that the cooperative motion system facilitates or impairs the input feeding the static association field.

Abstract (Italiano)

Questa tesi investiga i meccanismi responsabili dell'integrazione di segnali locali (siano essi orientazioni, posizioni o direzioni di elementi locali) in configurazioni globali. Il lavoro si compone di tre studi, che provano a dare una risposta alla domanda attraverso l'utilizzo di un paradigma di integrazione di contorni. Ciascuno studio approfondisce uno specifico aspetto del problema.

Il primo studio confronta due modelli di integrazione: il primo è il celebre "campo associativo", basato su connessioni laterali (presenti nella corteccia visiva primaria) tra campi recettivi adiacenti e sensibili per orientazioni locali simili. Il secondo modello è un filtro di second'ordine che riceve come input il risultato di un processo di rettificazione dell'output filtri di primo ordine.

Il secondo studio verifica, invece, se un sistema cooperativo locale spiega in maniera esaustiva l'integrazione di segnali locali di direzione. Inoltre, questo studio investiga anche la possibilità che il suddetto meccanismo cooperativo possa spiegare la "facilitazione data dal movimento" che si trova, di solito, quando si confronta la abilità di rilevare un la presenza di un contorno dinamico rispetto ad uno statico. In ultimo, lo studio tre amplia i risultati del secondo studio, avvalendosi di una tecnica di registrazione di potenziali evocati elicitati da contorni statici e dinamici.

Nel complesso, i tre studi supportano l'idea che un sistema basato su connessioni laterali (presenti nella corteccia visiva primaria) possa determinare l'integrazione di contorni statici, mentre un sistema cooperativo spiega l'integrazione di segnali di movimento locali. In aggiunta, questi due sistemi interagiscono continuamente, con il sistema di movimento che determina la qualità dell'input che sarà utilizzato, successivamente, dal sistema associativo statico.

Introduction

Human beings are mainly *visual* animals. That means that we explore the world with our sight. And our sight allows us the almost immediate recognition of objects, faces and places. This happens via a fast processing that the visual system performs on visual scenes. Nonetheless, research in the past seventy years has shown that this process, although really effective, is indeed far from being simple.
According to a rather simplistic description, the image that is projected on the retina is extremely fragmented. A series of contrast borders describe what we see and need to be decoded and integrated. The system is therefore asked to group together elements belonging to the same image and to exclude those belonging to other images. The result of this operation is important, because it will feed higher stages of visual processing till object recognition. It is not surprising, then, that a huge branch of research in vision has focused on these first stages of visual integration.

A powerful way to investigate integration is surely given by studies on *contour detection*. In these studies, observers are usually asked to detect an open contour-shape embedded in a field of distracters. Thanks to this paradigm, Field and col. (1993) proposed a rule to predict contour integration, i.e. to be grouped in a contour, elements have to respect a joined constraint of position and orientation. This constraint is the fundament of Field and col.'s successful model of contour integration: the association field model. Since its publication, however, a number of studies have come in succession to test its predictions. Some of them have tried to found neural correlates of the association field (sometimes coming to different conclusions; Li & Gilbert, 2002; Shpaner et al., 2013). Others have focused on the properties of this model, e.g. tuning to physical attributes of the stimuli etc. Others have questioned the fundamental rule itself, suggesting that the association field might be "overconstrained" (Watt et al., 2008).

There are still open questions. One of them refers to the mechanism responsible of integrating static disconnected elements into contours. Indeed, two models still compete: one is the association field proposed by Field, Hayes, and Hess (1993). The

other is a second-order filter, output of a rectification process of first-order filters, unsensitive to local properties.
In the first study (Chapter 3) I will explicitely try to disentangle between the two models of contour integration. In addition, I will test the role of the orientation cue only for two straight contour types (snakes, formed by elements parallel to the contour axis and ladders, formed by elements perpendicular to the contour itself). Moreover, I will investigate some low-level features of the stimuli (i.e. I will manipulate phase and spatial frequency) to assess whether integration involves early visual processing, that are selective to spatial frequency and phase. If contour detection performance is modulated by background and phase manipulation, then a mechanism based on lateral interactions is more likely to account for these results.

Another open question is how static local signals interact with dynamic signals. Indeed, recent studies are revealing that motion and static information are not independently processed, but constantly interact (Mather, Pavan, Bellacosa Marotti, Campana, & Casco, 2013).
In the second study (Chapter 4) I will investigate whether the extraction of local motion signals interact with the integration of local orientation signals. Previous studies have pointed out that form and motion processing interact, but the way such interaction takes place is not clear yet. Specifically, I will test whether motion extraction intervenes either at a local level (facilitating the grouping of locally moving signals) or after contour integration (by enhancing the saliency of the whole structure).

In the third study I will further investigate what emerges from the second study, i.e. whether grouping by motion occurs prior to the extraction of the static configuration. The main purpose of this study is to differentiate the dynamics of static- and dynamic-contour integration and to point out any difference existing in the analysis of different contour types. The psychophysical contour integration paradigm will be integrated with VEPs recording on the scalp.

1.1. Outline

This manuscript is organised in five chapters. Chapter 2 will provide an exhaustive review of the literature. It will summarise contributes of previous researches about contour integration and motion integration mechanism. The last paragraph of this Chapter will focus on studies that have investigated the electrophysiological correlates of both contour integration and motion.

Chapter 3 describes a study about the association field tuning to low-level properties of the stimuli. Chapter 4 contains an investigation of the mechanism responsible of the integration of moving contours, with reference to the classic association field model. Finally, in Chapter 5 the electrophysiological correlates of static and dynamic contour integration will be presented. For all the studies goals, method and results will be presented in detail and discussed.

Chapter 6 will provide a general discussion: findings in the manuscript will be interrelated and explained in accordance to an association field model based on lateral interactions between similarly oriented adjacent receptive fields.

Chapter 2. Previous studies about contour integration and motion perception

The world outside is extremely complex: objects have several features and colours and occlude each other. Sometimes they move, sometimes they are camouflaged (such as a tiger in the jungle), sometimes they lie in the dark. In same cases, things that were invisible while static become visible when they start to move (like a fish hiding in the sand). Moreover, the information that passes through the eyes does not carry all these details at once: it is fragmented and impoverished. And still we are able to correctly perceive and recognise objects without (much) effort.

Given the importance of a task like *perceiving things*, it is not surprising that a wide number of studies have investigated how our brain, and specifically our visual system, performs it.

In this chapter, I will review some of these studies and I will especially focus on those that investigated early stages of visual analysis, i.e. where local signals are bound together. I will cover a brief history of the topic and summarise the recent state of the art. Since studies addressing the topic come from different disciplines, this review will separately focus on contributions from psychophysics, neuroscience and computational science.

2.1. From perceptual grouping to contour integration

When an object is presented in isolation, it is easily segmented from the environment. See for example Figure 1b: Tigger is immediately spotted and recognised. What happens when the same object is no longer presented alone? In Figure 1a we can see what a challenge for the visual system segmentation can be. Tigger is now embedded in a very dense field of other objects and nonetheless we are still able to perceive its boundaries.

It becomes clear that early stages in visual analysis need to accomplish two main tasks: i) identifying the boundaries of an object ii) pooling together elements belonging to the object from large areas on the visual field and classifying the others as distracters. In other words, mechanisms that integrate local information into a global object representation must combine only the signals corresponding to single objects but keep

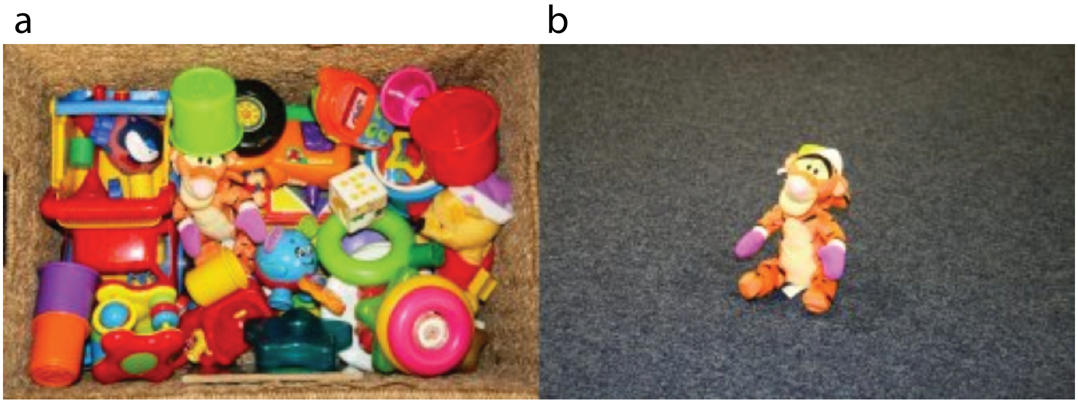


Figure 1. Segmentation problem. When an object as this Tigger puppet is presented in isolation (b) its boundaries are easily distinguished from the background. The same puppet in a crowded field (a) is still detected but its contours are now partially occluded by the other toys in the box. The visual system is asked to fill the occlusion gaps and guess the hidden shapes. (Taken from University of Ulster, http://biomed.science.ulster.ac.uk/vision/-Crowding,52-.html)

separate those that belong to different objects. These two tasks can be considered as two sides of the same coin: in both indeed the visual system is asked to *group* things together (either elements on a contour or features belonging to the same object). Perceptual grouping can then be defined as the phenomenon for which some elements in the visual field are perceived "going together" more than others (Wagemans, 2012).

Perceptual grouping is therefore a fundamental aspect of visual perception: it is what determines the emergence of a structure in perceptual experience. For this reason, it has been extensively studied since the early twentieth-century. Gestalt psychologists ((Max. Wertheimer, 1923; Max Wertheimer, 1938); see (Wagemans, 2012)) first attempted at elaborating *rules* of perceptual grouping. Max. Wertheimer (1923) came with a series of principles that allowed to systematically predict when grouping could occur.

Figure 2 gives an overview. The principles were proximity, similarity, uniform density, common fate, direction, good continuation and "whole properties" (closure, equilibrium, symmetry). To summarise, grouping emerges when the elements involved *share* some feature: they might be perceptually similar, go to the same direction or lead to the so-called "good continuation". The present manuscript will especially focus on "Common fate" (in Fig. 2f, all else being equal, elements that move in the same way tend to be grouped together, (Max. Wertheimer, 1923; Max Wertheimer, 1938)) and "Good continuation" (observers perceive Fig. 2g as containing two continuous intersecting lines rather than as two angles whose vertices meet at a point) (Max. Wertheimer, 1923; Max Wertheimer, 1938).

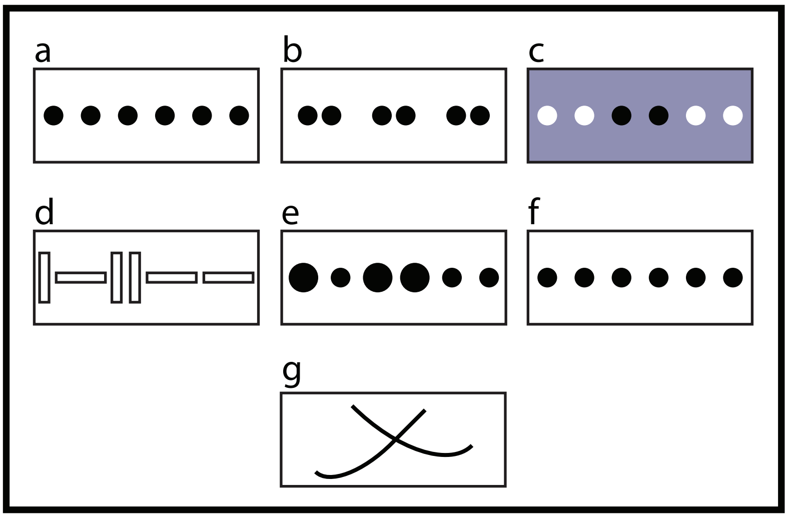


Figure 2. Gestalt grouping principles as proposed by Max. Wertheimer (1923). The Figure shows (a) no grouping and grouping by (b) proximity (c) colour similarity (d) orientation similarity (e) size similarity (d) common fate (g) good continuation. Figure adapted from Wagemans (2012).

The aim of the Gestalt psychologists was not so far from the focus of modern research on perceptual grouping: they were interested in how a global structure emerges, in other words how we perceive things as a "whole" (Wagemans, 2012).

These principles, although descriptive and sometimes simplistic, have been recently applied to explain how the visual system may solve the segmentation problem illustrated above.

Indeed, when presented with a natural scene, object boundaries need to be detected. Since contours are often fragmented and occluded by nearby objects, also segments composing the contour need to be grouped. In particular, proximity, similarity and good continuation are thought to play a special role in contour integration mechanisms (Wagemans, 2012).

To summarise, investigating contour integration gives us the opportunity to study pooling of local signals into a whole and have a deeper understanding of the way visual system works. More importantly, it has repeatedly been showed that the contours of a shape contain a huge amount of information, sufficient, e.g., to recognise objects (Attneave, 1954). Even fragmented contours can be used to infer what shape has been presented (Biederman, 1987; Kennedy & Domander, 1985; Panis, De Winter, Vandekerckhove, & Wagemans, 2008; Panis & Wagemans, 2009). The following sections will describe contour integration studies into details.

2.1.1. Contour Integration: Psychophysics

Although contour integration has always been a central topic in vision research, the interest in it has largely increased since Field and col.'s seminal paper on the so-called "association field" (Field et al., 1993). The authors moved again from the assumption that contours in natural scenes are often fragmented. Moreover, highly fragmentation increases the probability that boundaries of different objects are bound together by mistake. Thus, the visual system needs to develop a strategy to i) integrate fragments belonging to the same contour, ii) suppressing the distracters, i.e. fragments of other contours. To directly investigate this strategy, they developed a new paradigm for the study of contour integration. They used contours made up of Gabor patches and embedded them in a field of randomly positioned and oriented Gabor distracters. Observers were presented (in a 2AFC paradigm) with two displays in random order, one of which was composed by distracters only, whereas the other contained both contour target and distracters (Figure 3a). Observers' task was simply to press a key to indicate which of the two intervals contained the target. Performance in this task resulted extremely good, provided some geometrical constraints within the contour. Field et al. found that segmentation was particularly accurate when Gabors in the contour followed a joint constraint of local orientation and position, i.e. elements that are collinear and lie along the contour backbone are more easily grouped together (note that the joint constraint might be considered as an example of grouping by good continuation (Wagemans, 2012)). Jittering either their position or orientation led indeed to poorer performance (Field et al., 1993). Moreover, Field and col. also tested detection for different contour curvatures and found that performance dropped at orientation differences of up to ±60° amongst elements. Increasing the number of elements improved performance instead.

The authors introduced the notion of ''association field" to describe the geometric relationships between neurons that result in binding (Figure 3b). According to this idea, if two filters (approximated by two Gabors) are associated, they are linked and segregated from the background. The association is strong along the axis given by cell orientation but is not restricted to this axis. Indeed, connections are broad-tuned and can explain integration for large curvatures.

This paradigm is undoubtedly a powerful cue to study shape perception. Indeed, the emphasis here is on the extraction of a global structure, rather than on the perception of single local elements (Loffler, 2008). Moreover, to detect a contour it is necessary to integrate local information arising from large parts of the visual field, that cannot be covered by a single receptive field at the time (R. Hess & Field, 1999). We will return to this point in Section 2.1.2.

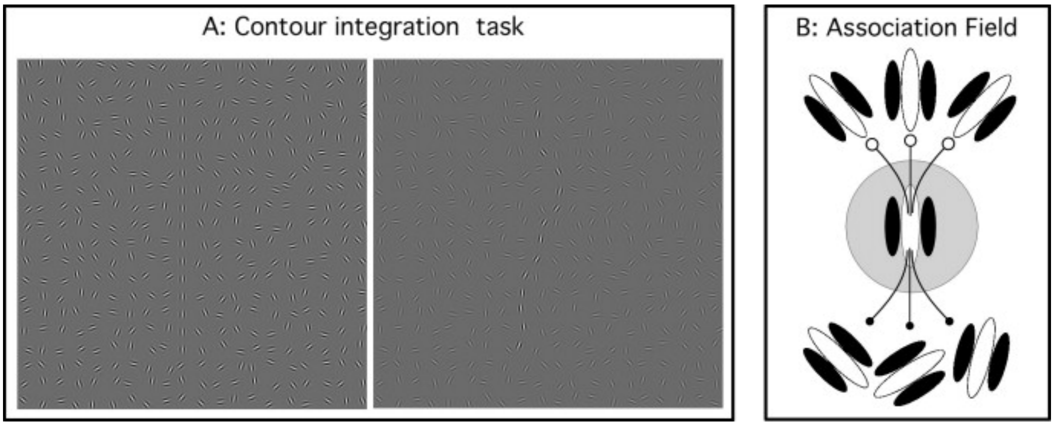


Figure 3. (a) The contour integration paradigm Left panel shows a straight contour made up of aligned Gabors and embedded in a field of similar randomly oriented and positioned patches. Right panel shows a curve contour. For explicative reasons, the contrast of the elements in the background has been halved. (b) Schematic representation of the association field model. The central Gabor can be successfully linked to the elements that follow a joint constraint of position and orientation (empty circles, top); binding is instead disrupted by elements that do not respect the orientation constraint (solid circles, bottom). The figure was published in Loffler (2008).

The properties of the association field are known. Field, Hayes, and Hess (2000) tested its sensitivity to changes in contrast polarity of the elements in the contour. In their previous paper (Field et al., 1993) they had randomised the contrast polarity of both the target and the distracters and found that detection remained unchanged. However, some criticisms arose: indeed, the Gabors used were narrow-band and thus not ideal to study a phase-selective mechanism. Moreover, randomization led to a mean difference of 90° in phase polarity between adjacent elements. They overcame these limits by using both even and odd patches that had either same or alternate phase. They found lower sensitivities when using alternate-phase targets, even if performance remained above chance and thus concluded that a mechanism both sensitive and insensitive to phase might be responsible for a flexible integration of contours.

An effect of alternating phase was also found by R. F. Hess and Dakin (1997) who reported a drastic loss in sensitivity when contours had alternate phase and were presented in periphery. The authors suggested that contour integration might be a primarily foveal function (R. F. Hess & Dakin, 1997, 1999), whereas a filtering mechanism might explain contour detection in periphery. Nevertheless, Nugent, Keswani, Woods, and Peli (2003) showed that the rate of reduction with eccentricity was more limited and similar to that of visual acuity. An alternative explanation has recently been given by May and Hess (2007b), who found that the eccentricity effect was only present for contours that did not follow the joint constraint of position and orientation. The association field, in their opinion, acts like an integration field (Pelli, Palomares, & Majaj, 2004) that pools together elements over a space that increases with eccentricity. Thus elements not following the joint constraint of position and orientation had weaker linkage and this caused a larger crowding effect exerted by distracters on target elements.

Also the association field tuning to spatial frequency has been investigated. Dakin and Hess (1998) measured observers' sensitivity to contours whose elements had alternate spatial frequency. The tolerable difference in frequency between elements was found to be dependent upon element orientation. Indeed, observers could tolerate 1.3 octave difference in frequency when the target was straight. Tolerance decreased to 0.7 octaves for curved paths. In another paper (Dakin & Hess, 1999) they tested the tuning of contour integration to spatial scale. They alternated broadband and narrowband elements to form a contour, the rationale being that a scale invariant mechanism would anyway integrate contours. However, this was not the case: contours consisting of alternating Gabor and step-edge micro-patterns were hard to detect, suggesting that contour integration does not occur independently within separate spatial frequency channels. To summarize, contour integration is generally limited within a spatial scale (being tuned for intermediate frequency level and disrupted by alternating the spatial scale) (Dakin & Hess, 1998, 1999), but is possible when elements differ in bandwidth (Dakin & Hess, 1999).

Other studies have showed that differences in disparity appear to have little effect on integration (R. F. Hess & Field, 1995; R. F. Hess, Hayes, & Kingdom, 1997; Kovacs, Papathomas, Yang, & Feher, 1996).

Special attention has been given to the local orientation constraint. Bex, Simmers, and Dakin (2001) introduced a distinction between contours formed by collinear elements (*snakes*) and contours formed by elements perpendicular to the contour backbone (*ladders*). Whereas Field et al. (1993) claimed a drastic impairment in performance when jittering the local orientation of the elements in the contour, Bex et al. (2001) found that performance decreased when observers were asked to detect ladders. However, it was far from being at chance. Similarly, Ledgeway, Hess, and Geisler (2005) tested integration for different local orientation values and, surprisingly, found detection to follow a u-shaped curve: the better performance was achieved with snakes and ladders, whereas contours formed by 45° oriented elements were the harder to be spotted (see also (Ledgeway & Hess, 2006)). Two classes of models might explain these findings. The first assumes that two different mechanisms account for the integration of snakes and ladders, i.e. an association field for snakes and second-order filtering for ladders, as suggested by models of texture segregation (N. Graham, Beck, & Sutter, 1992; N. Graham & Sutter, 1998; N. Graham, Sutter, Venkatesan, & Humaran, 1992; Lin & Wilson, 1996). However, other studies have pointed out that it is unlikely that different mechanisms underpin the integration of the two different contours (May & Hess, 2007a). Another possibility is that the joint constraint of position and orientation needs to be relaxed and perhaps to predict weaker binding within adjacent elements oriented orthogonally to the contour path (Ledgeway et al., 2005; May & Hess, 2007b; Yen & Finkel, 1998).

Hence these results have challenged the assumption of an association field, together with other studies claiming that position and inter-element spacing are the most crucial cue in contour integration (R. Watt, Ledgeway, & Dakin, 2008). It has even been suggested that the association field might be over-constrained (R. Watt et al., 2008). The study in Chapter 3 will accurately address this question.

Finally, some authors have suggested that the association field model should also take into account background modulation effects. Dakin and Baruch (2009) varied the orientation features of the background elements surrounding the target. They tested participants in a shape discrimination task using both snakes and ladders. Better performance was obtained when surrounding elements were near-perpendicular to the elements in the contour. Conversely, poorer performance was caused by near-parallel surrounding elements. This was true for both snakes and ladders. These results showed a suppressive surround modulation that was further confirmed with a simple contour detection paradigm (Robol, Casco, & Dakin, 2012) and with straight contours as target (Schumacher, Quinn, & Olman, 2011). A possible explanation was a crowding-induced local orientation uncertainty (i.e., contour elements are grouped with distracters), in line with other works showing a strong connection between crowding and association field models (Chakravarthi & Pelli, 2011; May & Hess, 2007b).

2.1.2. Contour Integration: Neural Substrate

To fully answer the question of how the visual system binds local signals into contours, it is essential to take a look at the underpinning neural machinery. This section will thus address the role of visual area V1. It will also discuss the criticism on the role played by this area in contour integration.

A large number of physiological studies have brought evidence that visual area V1 is mainly responsible for contour integration. Primary visual cortex has indeed a columnar organization specialised in responding to different orientations at different spatial scales, direction and ocular dominance (Hubel & Wiesel, 1977). In their review, R. F. Hess, Hayes, and Field (2003) hypothesised that such a structure might facilitate communication between cells that respond to similar properties but are dislocated in different parts of the visual field. In addition, long-range lateral links between cells with similar orientation tuning have been found in area V1 of different animals (Gilbert & Wiesel, 1989; Rockland, Lund, & Humphrey, 1982; Schmidt, Goebel, Lowel, & Singer, 1997). The fact that cells in V1 are band-pass in both orientation and spatial frequency is also consistent with psychophysical data on contour integration (Dakin & Hess, 1998, 1999). Also computational studies on statics of natural images are consistent with the idea that cells V1 anatomy is optimised to respond to contours: indeed, natural images mainly contain parallel and aligned structures (Geisler, Perry, Super, & Gallogly, 2001). The aligned structure follows the “co-circularity rule”, i.e. pairs of separated local contour parts are most likely to be aligned along a linear or curved path. This pattern occurs at different spatial scales (Sigman, Cecchi, Gilbert, & Magnasco, 2001). The co-aligned information is due to the contour structure in natural images. The parallel information, on the other hand, is most frequently derived from regions of the same object or texture (Geisler et al., 2001). Consistently long-range interactions in V1 mainly occur among cells (with similar orientation preference) that fall on a co-circular space oriented along the axis of the cell's orientation preference, matching the contour statistics of natural images (Bosking, Zhang, Schofield, & Fitzpatrick, 1997; Gilbert & Wiesel, 1989; Malach, Amir, Harel, & Grinvald, 1993).

Psychophysical findings by W. Li and Gilbert (2002) are also in close agreement with the idea that long-range interactions are responsible for contour integration. The authors reasoned that, if binding occurs because of connections within V1, then the maximum extent over which elements are combined should not exceed the extent of the connections, i.e. they should have found a spatially limited extent of global contour integration. Their stimuli consisted of contours made up of straight non-overlapping lines embedded in a field of randomly oriented lines and varied contour length and inter-element spacing. They found that contour salience (measured in a detection task) increased with contour length and decreased with increasing inter-element separation up to a critical distance of approximately 2°. Beyond this critical distance no contour could be detected, irrespective of the number of elements in the contour. This value is consistent with the spatial extent of horizontal connections found in V1 (Kapadia, Ito, Gilbert, & Westheimer, 1995; Stettler, Das, Bennett, & Gilbert, 2002). In line with this, collinear facilitation also disappears when target and flanks are separated by more than 2° (Dresp & Grossberg, 1997; Kapadia et al., 1995), although it is strongest for much smaller separations. These findings were later replicated with monkeys (W. Li, Piech, & Gilbert, 2006). Changes in salience of the contours were highly correlated with the activation recorded from V1: more salient contours (obtained by increasing their length) were correlated to a monotonically increment in neuronal responses. Conversely, less salient contours (obtained by increasing inter-element spacing) corresponded to a decrement in neuronal responses.

More recently, Gilad, Meirovithz, and Slovin (2013) presented macaque monkeys with a contour detection task and retinotopically mapped both contour and background on monkeys' area V1 with an imaging technique. They found that early responses (40-140ms) were correlated to the presence of oriented elements in general, with neurons responding to both contour and no-contour displays. Later response (150-250ms) showed, instead, slightly higher activation in the areas were the contour was mapped and suppression of response in correspondence of background elements. They concluded that late phase responses might be correlated to grouping, rather than to specific stimulus features. Moreover, their findings suggest that grouping is achieved especially through background suppression.

It is worth noting, however, that a number of studies have recently questioned the prominent role of primary visual cortex in favour of a higher-level integration process (Samonds, Zhou, Bernard, & Bonds, 2006; Shpaner, Molholm, Forde, & Foxe, 2013). According to this approach, inputs from primary visual cortex converge on cells with larger receptive fields in higher-order stages of visual processing, e.g. the lateral occipital complex (Hubel & Wiesel, 1962; Spillmann & Werner, 1996). An attempt to disentangle the issue has recently been done by Shpaner et al. (2013). In this study ERPs were recorded in a contour detection task. The stimuli could be displayed in a quadrant portion of the visual field at the time. Such a presentation is useful to elicit an early component (C1, 50-100 ms), usually associated to activation in area V1. The authors found no modulation on this early component when either a contour or background-only displays were presented, suggesting no involvement of V1 in contour segmentation. They instead found later modulation (N1, 160-200 ms), which they associated to activation in higher visual areas. Interestingly, they also reviewed previous works linking lateral connections and contour integration, showing that activation in primary visual cortex in response to global contours emerged later than response to local orientation signals. They considered these findings to mirror feedback signals to V1 and pool of local signals in higher cortical areas. Accordingly there is evidence of recurrent connections modulating activation of area V1 (Di Russo, Martinez, & Hillyard, 2003; Lamme & Spekreijse, 2000).

To conclude, it is not fully clear at what stage of visual analysis contour integration takes place. There is evidence in favour of a local processing carried out by neurons in V1 as well as a more global and higher-level process that causes activation in V1 via feedback connections. The study in Chapter 5 will investigate the topic more in detail by using electrophysiological techniques.

2.1.3. Contour Integration: Computational Models

The association field model is surely the most successful model of contour integration. Nonetheless, recent psychophysical data have challenged its assumption of a joint constraint of position and orientation. Moreover, also contextual organization seems to exert an effect on contour integration (see above). To account for all these results, other models have been proposed. On one hand, they give alternative interpretation of the phenomenon. On the other hand, they aim at extending and improving the association field model. In this section I will shortly review these two classes of models.

An interesting model of contour integration entirely based on V1-like structure was proposed by Z. Li (1998). The model is a neural network that resembles area V1, with its recurrently connected excitatory neurons, inhibitory interneurons and oriented receptive fields. The model could replicate data from psychophysical experiments, such as length, smoothness and closure. Moreover, since it incorporates both surround orientation inhibition and contour linking, can account for a range of pop-out effects. This and similar models (Z. Li, 2000; Zhaoping & May, 2007) are thus useful to predict background effects as in (Dakin & Baruch, 2009).

However, this model does not provide connections between parallel, rather than collinear, receptive fields. To this extent, it is not far from the association field model that does not provide connections between elements forming contours not fully respecting the joint constraint of position and orientation. Nonetheless, detection of these contours is often possible (as in the case of ladders) (Bex et al., 2001; Ledgeway et al., 2005). An extension of Field's model (Field et al., 1993) proposed by Yen and Finkel (1998) explicitly includes excitatory, long-range, horizontal connections both between adjacent oriented receptive fields that sample regions of space (i) along smoothly varying curves ("co-axial") and also (ii) those that have the same preferred orientation but are orthogonal to those curves ("trans-axial"). In other words, connections for both snakes ("co-axial") and ladders ("trans-axial") are provided. Moreover, excitatory connections in this model are limited to these two configurations and are absent between receptive fields that have 45° preferred orientations relative to a contour, thus explaining poor performance with these types of contours found by (Ledgeway & Hess, 2006; Ledgeway et al., 2005). Furthermore the trans-axial connections are more spatially focused, with the degree of excitation falling off much more rapidly with distance than for the co-axial connections. However it has been underlined (Ledgeway et al., 2005) that the relatively good performance found when elements are orthogonal (relative orientation 90°) to the contour they represent may pose difficulties for computational schemes that suggest that inhibitory, rather than excitatory, trans-axial connections exist between neighbouring receptive fields (e.g., (Z. Li, 1998; Ursino & La Cara, 2004)).

Since this model compares activation between configurations and switches off the less-active path, it directly allows competition between locally parallel and perpendicular configurations (as in the paradigm tested by (Dakin & Baruch, 2009; Robol et al., 2012)). This architecture could explain the disruptive effect of parallel surrounds on snakes (which would maximize the probability of the co-axial/incorrect configurations winning) and the facilitatory effect of perpendicular surrounds (which would minimize the probability of interference from the surround), this way potentially accounting for contextual effects found by (Dakin & Baruch, 2009; Robol et al., 2012).

Models described so far have two main features: they (i) are based upon good continuation principles and (ii) rely primarily upon cortical networks in area V1 and V2 (Z. Li, 1998; Yen & Finkel, 1998). Nonetheless, studies on contour integration and imaging methods are increasingly showing the involvement of areas such as V4 (Pasupathy & Connor, 1999) and LOC (Shpaner et al., 2013) in contour integration. Moreover, these models fail at incorporating global-shapes effects (e.g. high salient spirals that are not seen by models that code for closed shapes). Anyway, as argued by Loffler (2008), the association-field itself was never intended to capture such global effects. It is undoubtedly successful describing grouping between elements, however growing body of evidence indicates that other, more complex and more global computations are also involved in contour integration.

R. Watt et al. (2008) recognised the limitations of the association field model and attempted to predict integration of local signals into contours using other cues than the joint constraint of position and orientation. They moved from two main considerations: (i) density of elements influences performance in a contour detection task (Kovacs & Julesz, 1993; W. Li & Gilbert, 2002); (ii) detecting a contour is a different process from recognising the features of a contour.

They described performance in a contour detection task as the maximum tolerable contour curvature as a function of path length. They then compared these data with predictions from the classical association field and two models based only on spatial adjacency (i.e. positional cue). Their models differed in the sense that the first one simply pooled together elements because of a "where" instead that a "what" cue (isotropic linking), according to a critical distance. The second pooled together elements again regardless of their orientation (anisotropic linking), but with the constraint that one element could be linked to any others nearby along the orientation of its axis. They interestingly found that all the models could account for their data, thus binding simply based on "where" (rather than "what") seems enough to explain contour detection.

Different from the association field are models predicting integration by means of a filtering process. Specifically, this process has been borrowed from studies on texture segmentation (N. Graham, Beck, et al., 1992; N. Graham & Sutter, 1998; N. Graham, Sutter, et al., 1992) and accounts for detection of ladder contours. In these models, a small-scale linear filter is first followed by a nonlinearity (e.g. squaring, full-wave rectification), and then by a large-scale linear filter with orientation orthogonal to that of the first filter. Such a mechanism gives a strong response at a texture border (such as a border between two areas of different orientation) but a weak response elsewhere. It should predict detection of fairly straight ladders: a small-scale filter aligned with the elements will give strong positive and negative responses along the path of the contour; if these responses are rectified, then a large-scale filter orthogonal to the small-scale filter will give a positive response along the whole of the contour. However, data from May and Hess (2007b) question this assumption: indeed, such a model could not predict failure in ladder detection at increasing eccentricities. We will come back to this issue in the Discussion section of the following chapter.

2.2. Motion Perception and Perceptual Grouping

The introduction so far has focused on grouping of static local signals. I have outlined what happens to single segments before they become part of an object. Sometimes, however, the visual system needs to pool moving signals together to extract a shape, as it happens, for example, when a fish that was invisibly hiding in the sand starts to move and becomes visible. In this section I will address the unresolved problem of the interaction between shape- and motion-perception. Additionally psychophysical evidence and neural basis of grouping by motion will be presented. A special attention will be dedicated to previous literature on motion and contour integration.

2.2.1. Neural basis of motion perception

We have delineated the neural regions involved in static processing (see par. 2.1.2). It is reasonable asking, then, what regions are instead involved in motion processing. Also in this case knowledge comes from electrophysiological investigations on animals and neuroimaging on humans. Traditionally MT (for a complete review on this area see: (Born & Bradley, 2005)) and MST have been the most studied motion-selective areas in animals. More recently, non-invasive neuroimaging techniques have revealed the presence of direction selectivity areas in human observers as well. They have confirmed the important role of MT and MST and also revealed a huge number of areas involved in motion processing (e.g. V3 (Tootell et al., 1997), KO (Van Oostende, Sunaert, Van Hecke, Marchal, & Orban, 1997)), suggesting that "the entire visual brain [could be] a motion area" (Culham, He, Dukelow, & Verstraten, 2001).

The basic idea is that local motion is processed in V1, while MT is more specialized in processing global patterns of motion. This is both related to structural and functional properties of these areas.

There are indeed direction-sensitive neurons in V1 (Hubel & Wiesel, 1968) The simple extraction of local motion signals, however, leads to the so-called "aperture problem", i.e. moving edges seen through small apertures appear to move orthogonally to their length. Indeed direction-selective neurons in V1 extract orthogonal velocities (Hubel & Wiesel, 1968; Movshon, Adelson, Gizzi, & Newsome, 1985). When motion is analysed at high resolution, therefore, different parts of the object will be extracted by small, orientation-dependent filters. Each sampled part of the object will thus appear to move in different directions, depending on the orientation of the edge in that position (Wallach, 1935). Neurons in V1 act like linear space-time filters (McLean & Palmer, 1989; Reid, Soodak, & Shapley, 1987); they measure the amount of energy (i.e. they *see* motion) but they don't know the stimulus veridical direction (Adelson & Bergen, 1985).

Direction needs to be computed somewhere else, for example in area MT (See Paragraph 2.2.3 for rules of integrating local motion signals).

MT exhibits a wide percentage of neurons having direction selectivity (Lagae, Maes, Raiguel, Xiao, & Orban, 1994). Two segregated populations of neurons can be distinguished (Born & Tootell, 1992): one has a facilitatory surround, which, if stimulated, enhances response to a central stimulus. The other (at least half of the total number of neurons) has instead an antagonistic surround, i.e. stimuli in the surround inhibit response to a centrally presented stimulus (Allman, Miezin, & McGuinness, 1985; Bradley & Andersen, 1998; Raiguel, Van Hulle, Xiao, Marcar, & Orban, 1995). These structural properties lead to functional specialization: the former population might sum motion cues coming from relatively large regions of the visual field, thus processing global motion (Born & Tootell, 1992). Summation is larger at low contrast (Pack, Hunter, & Born, 2005). The latter might be responsible for analysis of local motion, helping the extraction of discontinuities in motion direction. Consistently with this second function, MT seems also to have a prominent role in reducing noise in motion processing: it shows indeed *motion opponency*, that is, neurons responding to dot patterns moving in their preferred direction are suppressed by simultaneously introduced dot patterns moving in non preferred directions (Snowden, Treue, Erickson, & Andersen, 1991). Opponency is released when differently moving stimuli are presented at different disparities: as a general rule, the more a stimulus sticks out in terms of direction and depth, the larger the neuron's response will be (Bradley & Andersen, 1998). These findings support the idea that MT neurons are critically involved in segmenting an image into separately moving parts (Born & Bradley, 2005).

Studies on monkey's visual area MST indicate neurons that encode the global patterns of motion (optic flow) usually created by forward locomotion through the environment (Graziano, Andersen, & Snowden, 1994; Tanaka & Saito, 1989). At least 25% to 35% of neurons present in the dorsal part of the area MST (i.e., MSTd) in the macaque visual system have large receptive fields (from 10° up to 100°; (Desimone & Ungerleider, 1986; Tanaka & Saito, 1989)) and show selectivity to optic flow and to its components (Duffy & Wurtz, 1991; Graziano et al., 1994; Lagae, Raiguel, & Orban, 1993; Tanaka, Fukada, & Saito, 1989).

Studies on humans have revealed more or less the same pattern of findings. The human homolog of monkey areas MT, MST is believed to be V5, an area adjacent motion selective cortex (Zeki et al., 1991). This region can be parted in at least two subcomponents: one, more posterior, corresponding to MT and the other, more anterior, to MST (Dukelow et al., 2001; Huk, Dougherty, & Heeger, 2002). Thus it is often referred to as MT+ or complex. This complex is one of the most easily and consistently activated regions in human cortex, responding to a wide range of dynamic stimuli (Culham et al., 2001). However, the features of the stimuli determine which subpart will be activated. Morrone et al. (2000) compared BOLD signal within the region to translational and optic flow stimuli, such as rotation or expansion. They found optic flow to activate a region that was ventral to the one that, instead, responded to translation. Similar results were also found by Pitzalis et al. (2013).

In an fMRI investigation, Tootell et al. (1995) tested tuning to motion-related properties in human area MT. They found higher contrast sensitivity in MT than in V1. MT activity was also lowest (though still present) for chromatic stimuli when isoluminant. Finally, the area was strongly driven by motion in general, e.g. by moving gratings, moving dots or incoherent flicker. These results are in agreement to the fact that MT mainly receives inputs from the magnocellular stream. Indeed, there is evidence that in areas receiving inputs from this stream there are progressively more direction-selective cells (Felleman & Van Essen, 1987; Gegenfurtner, Kiper, & Fenstemaker, 1996; Gegenfurtner, Kiper, & Levitt, 1997). Furthermore, this lack of motion-type specialisation has made researchers think that MT might have the role to pool together different types of motion signals (Born & Bradley, 2005).

Another area that is often related to motion processing is human area V3 (Tootell et al., 1997). Literature on this region shows that it is specialised to respond to second-order motion (Smith, Greenlee, Singh, Kraemer, & Hennig, 1998).

Finally, to the purpose of this review, it is worth mentioning that motion-defined contours (i.e. texture borders defined by changes in direction of moving dots) selectively activate the so-called *kinetic occipital cortex* (KO) (Dupont et al., 1997; Orban et al., 1995; Van Oostende et al., 1997). Van Oostende et al. (1997) compared BOLD responses to moving boundaries, luminance borders, uniform and transparent motion in KO and MT. KO was found to respond to kinetic borders only, whereas MT complex responded to motion in general, regardless of the specific stimulus display.

2.2.2. Motion perception and motion-form interaction

In the outside world objects are not always static and human beings are good at detecting moving objects, even if they move fast and we are not directly paying attention to them (e.g. a car that is approaching while we cross the road). Studies (Banton & Bertenthal, 1996; Wattam-Bell, 1994) have confirmed these outstanding capacities when they showed that human observers are capable of discriminating the direction of a certain number of dots embedded in a background of randomly moving distracter dots even when coherence thresholds (i.e. the number of coherently moving dots divided by the total number of dots) is as low as 5-11%.

Classically, motion and static perception have been addressed as separate and independent processes. Goodale and Milner (1992) in 1992 proposed the so-called "two-stream hypothesis", according to which two separate pathways are thought to especially encode for identity ("what" or *ventral* pathway) and spatial position ("where" or *dorsal* pathway) of objects. These two pathways are both functionally and neuroanatomically segregated (Mishkin & Ungerleider, 1982; Mishkin, Ungerleider, & Macko, 1983): the *what* pathway stretches from the primary visual cortex in the occipital lobe and forward into the parietal lobe. The *where* pathway also starts from the striate cortex, but continues into the temporal lobe. Inputs from the *parvo*-cellular system converge in the ventral system, whereas inputs from the *magno*-cellular system mainly converge in the dorsal system (Maunsell, Nealey, & DePriest, 1990).

Such segregation arose from studies on monkeys that demonstrated a double-dissociation when disconnecting the striate cortex from either inferior-temporal or parieto-occipital cortex: in the former, a pattern discrimination task was mainly impaired, whereas in the latter the monkey could not perform a visual-spatial task anymore (Mishkin & Ungerleider, 1982). Further studies demonstrated that the two streams are involved in different processes: neurons in areas belonging to the ventral stream show selectivity for colour, shape and texture while those belonging to the dorsal stream show selectivity for direction and speed of visual motion (Maunsell & Newsome, 1987; Ungerleider & Pasternak, 2004).

However, several studies have lately questioned the segregation hypothesis. A constant motion-form interaction is now believed to take place in the brain (for a review, see (Mather et al., 2013)). The principle of "common fate" postulated by Gestalt psychologists (Max Wertheimer, 1938) represents a good example of such interplay: according to this principle, an invisible form composed of randomly arranged dots against a distracters background becomes immediately visible as soon as it moves, by virtue of the "common fate" of its dots, all moving together and sharing common speed and direction (Edwards, 2009; Ledgeway & Hess, 2002; Uttal, Spillmann, Sturzel, & Sekuler, 2000). Thus, motion can help shape detection. Also, form can help motion extraction by providing, e.g., direction cues (Geisler, 1999). Besides, motion can alter a form cue, such as position, and make it shift towards motion direction, a phenomenon known as motion-induced position shift (De Valois & De Valois, 1991). Shape features, such as local orientation, can influence the ability to detect a coherent motion trajectory in a field of randomly moving elements (Pavan et al., 2011) and the ability to segment a motion-defined contour (Ledgeway & Hess, 2006; Ledgeway et al., 2005). Finally, the Lateral Occipital Complex (LOC), a ventral area involved in object perception, responds to objects regardless of whether they are defined by luminance, texture, or motion (Grill-Spector, Kushnir, Edelman, Itzchak, & Malach, 1998), suggesting that motion input does indeed reach the ventral stream.

Mather, Pavan, Bellacosa, and Casco (2012) explicitly aimed at showing motion-form interaction using an adaptation paradigm. The authors adapted to transparent motion (Verstraten, Fredericksen, & van de Grind, 1994) and found that the strength of motion adaptation (measured as motion aftereffect, MAE) was modulated by the presence of simultaneous orientation signals, i.e. oriented pedestal gratings. If the orientation of the superimposed pedestal during adaptation was orthogonal to the resulting unidirectional MAE, then the strength of the MAE was decreased relative to the condition in which the grating was parallel to the resulting MAE direction. Since transparent motion is thought to be integrated at the global motion level (MT, (Verstraten et al., 1994)), these findings provided evidence that form and motion interact at this level. In a follow-up study (Pavan, Marotti, & Mather, 2013), motion-form interaction was also tested beyond the global motion level, i.e. at the level were optic flow is extracted (MST, (Graziano et al., 1994)). The authors employed the same adaptation paradigm as in (Mather et al., 2012); they measured the duration of the phantom MAE (Weisstein, Maguire, & Berbaum, 1977) in which they adapted and tested different parts of the visual field, with orientation signals presented either in the adapting or non-adapting sectors. As before, they found that motion adaptation was suppressed most by orientation signals orthogonal to optic flow direction. They interpreted these results as evidence that motion-form interactions also take place at the global motion level, where optic flow is extracted.

To conclude, there is an increasing agreement on the idea that shape and motion perception are constantly interacting. Psychophysiological studies ((Thiele, Dobkins, & Albright, 2001); for a review (Rousselet, Thorpe, & Fabre-Thorpe, 2004)) are also showing the way in which interaction takes place in the brain. However, further investigation is certainly required. For this reason, in chapter 4 I will discuss and present data on the interaction between contours and motion inputs. Chapter 5 will also focus on the same topic combining psychophysics with event-related potential technique.

2.2.3. Pooling of local motion signals

As already pointed out, the receptive fields of early stages in the visual hierarchy are spatially limited and are highly selective for a limited range of stimulus attributes, including direction of motion (Anderson & Burr, 1987; Henry, Bishop, & Dreher, 1974; Schiller, Finlay, & Volman, 1976). The first stage of motion analysis is not a global estimate of the motion of an object, but rather a dense set of estimates of the motion present at each location in the visual field. These local estimates need thus to be grouped and combined to determine the motion of the object as a whole. Asking how pooling is performed can hence apply to local motion signals too. Investigating integration of motion-defined contours could represent a way to give this question an answer. Moreover, motion-defined contour integration and its tuning to both cues from static (local orientations, spatial frequencies) and dynamic (speed, temporal frequency) domains are a proxy to the study of motion-form interaction.

Nonetheless, to date only few studies have used dynamic contours as stimuli (note that motion in these experiments is usually implemented by making the carrier of the

Gabors drift in a direction perpendicular to their orientation). Bex et al. (2001) focused on speed tuning in a contour integration paradigm. Special attention was dedicated to assess the role of common direction: indeed, they used both snakes and ladders, because they both indicate regions of common organisation in the display. Moreover, Gabors could either drift (at the same vs. at different temporal frequencies) or counter-phase flicker. Ability was measured as orientation jitter threshold (i.e., the maximum tolerable amount of orientation jitter applied to local elements in the contour). Their findings revealed that snakes were always more easily detected than ladders. Furthermore, motion systematically helped detection, even when speed within the same contour spanned a range of three octaves. Also in-phase flicker increased tolerability to orientation jitter.

The interaction between local orientation and motion was also studied by Ledgeway et al. (2005), who confirmed that adding motion to orientation-defined contours helped detection and snakes were found to be generally more salient than ladders. They (Ledgeway & Hess, 2002; Ledgeway et al., 2005) also examined detection with motion-defined contours, i.e. the envelopes of the elements forming the contour enclosed a non-oriented 2D noise. The lack of orientation caused contours to be undetectable when static. The direction of contour patches determined performance in this condition: performance was better when elements drifted towards the next element on the contour (i.e., same speed, slightly different direction) than when they simply shared a "common fate". Consistently Verghese, McKee, and Grzywacz (2000) found detection thresholds for three aligned dots to be greater when they moved parallel to their mean orientation axis than perpendicular to it.

In another paper, Bex, Simmers, and Dakin (2003) examined how the spatial organisation of local directional signals determined the visibility of contours defined by motion. Observers' task was to detect a contour that could move across the entire display in a series of motion and contour conditions. They initially observed little difference between conditions in which the contour was moving rigidly (following a "common fate") and non-rigidly (each element was moving towards the next element on the contour). The two motion conditions did not differ also when speed, spatial frequency and spacing between elements were varied. As other global motion mechanisms reported elsewhere (Edwards & Badcock, 1994), the contour mechanism involved in this task was broadly-tuned to spatial frequency and not tuned to contrast polarity, since alternating these properties within the contour did not extremely affect contour visibility. The authors simply concluded that the visibility of moving contours was determined by the shape defined by the directions of the elements forming it.

Ledgeway and Hess (2006) also investigated sensitivity to spatial frequency, but using contours made up of oriented micropatches. Specifically, the aim was to find out whether mechanisms involved in static and dynamic contour integration had the same tuning to spatial frequency and orientation. Stimulus displays could be formed by either unoriented or oriented band-pass filtered elements, and this no longer constrained drifting to be perpendicular to micropatches' local orientation. Results showed again that spatial frequency tuning was relatively broad for both oriented and unoriented contours in the dynamic domain. Orientation tuning was instead found to be narrow: alternating the orientations of the micropatches in the contour largely worsened detection with both static and dynamic conditions, even when the elements had same spatial frequency and direction. Detection was generally better in the dynamic than in the static condition (most evident with 45° orientated elements). Again snakes were found to be more salient than ladders, regardless of whether they moved orthogonally or parallel to the contour axis. Finally, temporal modulation itself (e.g., flicker) did not enhance performance, further indicating the prominent role of direction cues.

To summarise, the mechanism underpinning the integration of moving contours seem to have its own peculiarities. Unlike the one responsible for static contours (Dakin & Hess, 1998), it can combine motion information across a range of spatial scales (Bex et al., 2003; Ledgeway & Hess, 2006). Similarly to static integration, instead, local orientation tuning is relatively narrowband (Ledgeway & Hess, 2006). Results so far suggest a prominent role of direction: when orientation cues are not available, directional signals are preferentially integrated along the trajectory of motion (Ledgeway & Hess, 2002). On the other hand, also a simple common fate allows detection (Bex et al., 2001, 2003; Ledgeway & Hess, 2002, 2006; Ledgeway et al., 2005). It is harder, instead, to describe the interaction between motion and local orientation signals. A static structure influences detection, with snakes always more salient than ladders (Bex et al., 2001; Ledgeway et al., 2005), regardless of motion direction (Ledgeway & Hess, 2006). However, motion cues help detection of contours in which orientation of the elements relative to the path is critical (Ledgeway & Hess, 2006).

Thus, although there are some similarities concerning the specific rules that allow the integration of local features into a global contour, importantly these rules are not identical for contours defined by motion signals and those defined by static orientation cues. However, the idea that the visual system supports a generic, cue-invariant association field is unlikely (R. F. Hess et al., 2003): two association fields for spatial and motion domains are characterised by different features that would rather suggest different underlying mechanisms. The studies in chapter 4 and 5 will try to disentangle this problem.

2.2.4. Models of local motion signals integration

As remarked above, it might exist a specific association field responsible for the integration of moving signals. It has been suggested that such a mechanism might follow grouping rules similar to those implemented in the static association field (R. F. Hess et al., 2003; Ledgeway et al., 2005). The way this mechanism should operate through is, however, still debated. Nonetheless, some solutions to the problem of pooling local motion signals have been proposed; in Figure 4 there is an overview. Some of them focus on the ability to compute a common direction, whereas others are related to the ability of the visual system to isolate coherently moving signal areas from the background.

Strategies for combining motion estimates were initially introduced to explain the neural computation of pattern motion in plaids, which are formed from the superposition of two 1-dimensional gratings (Adelson & Movshon, 1982). Moreover, integration of signals spanning different orientations is important to solve the aperture problem (Adelson & Movshon, 1982) and thus to extract a veridical direction of motion.

Classically, two different models compete to address this problem (for a review about MT role in solving this problem see (Born & Bradley, 2005)): the intersection of constraint (IOC; (Adelson & Movshon, 1982; Weiss, Simoncelli, & Adelson, 2002)) and the vector average (VA; (Kim & Wilson, 1993; Wilson, Ferrera, & Yo, 1992; Wilson & Kim, 1994)).

The former is composed by two serial stages (Figure 4a, red dot): motion signals are locally and independently analysed by oriented detectors selective to component directions and their velocities are extracted (stage one). The resultant velocities of the two components are then combined (stage two). Each component constrains the possible 2-dimensional pattern motion but a unique velocity could be chosen as the velocity that satisfies both constraints.

The VA (Figure 4a, green dot), instead, predicts the global perceived direction to be consistent with the vector average of the single components.

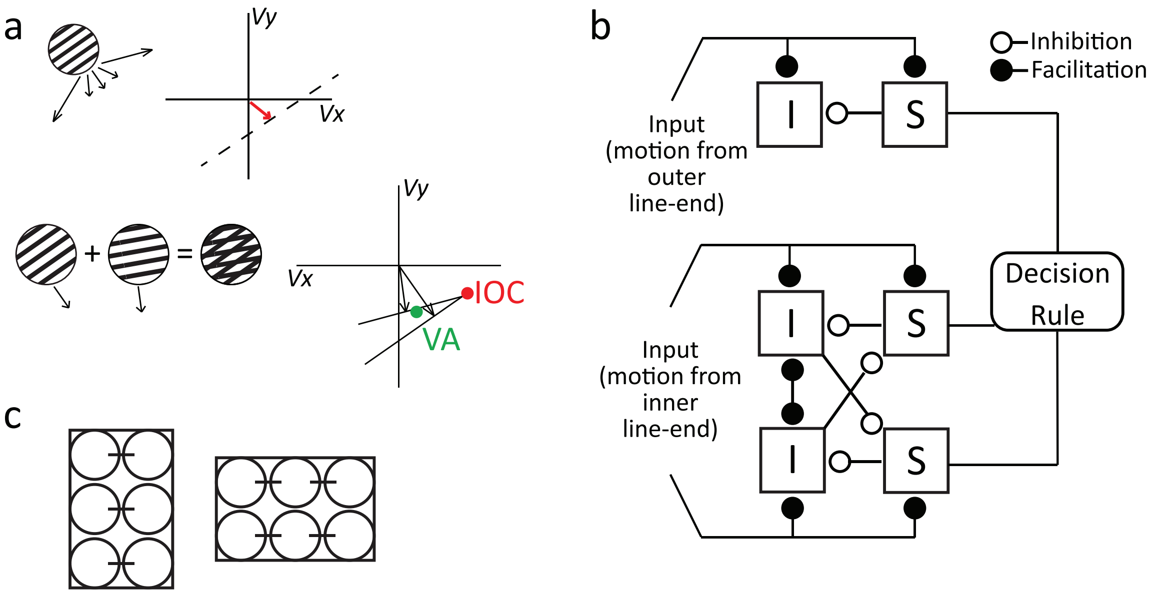


Figure 4. A schematic overview of models of motion pooling, as described in this paragraph. (a) Perceived direction of two superimposed gratings moving in different directions as predicted by the IOC (red dot) and the VA (green dot) rules. The diagram presented upper left describes the velocity space of a drifting grating; the vector represents the component normal velocity and the dashed line represents the global motion vector consistent with the normal component. According to the IOC rule, when two gratings are superimposed (below), they are first independently analysed and their separate velocity spaces are drawn. The velocity spaces generate two distinct constraint lines. The constraint lines for any two vectors lying on the same velocity space will intersect at a single point (second stage) representing the global motion (red dot). The VA rule suggests instead the perceived direction to be consistent with the average of the two components normal velocities (green dot). Adapted from (Weiss et al., 2002). (b) A cooperative-competitive system model involving long-range facilitation and surround suppression between collinear contours (Lorenceau, Giersch, & Series, 2005). The model comprises a population of "integration" cells (I), sensitive to collinearity and responding to stimuli outside the RF; a population of "segmentation cells" (S) that suppress activation from stimuli outside the RF; a decision rule based on the temporal integration of the activities of all S cells. The direction of the stimulus is discriminated when the decision unit reaches a given threshold. (c) Second-level facilitation in the direction of motion (Vreven & Verghese, 2002). The lines represent connections between circular low-level motion detectors tuned to the same direction and speed. When a detector is activated, it spreads activity in the direction of motion, in this case the horizontal connections indicate tuning to horizontal motion (that could be either leftward

of rightward). See text for further details.

Amano, Edwards, Badcock, and Nishida (2009) demonstrated that the perceived direction of a group of local 1D motion signals (e.g. drifting Gabors) was computed according to IOC rule, whereas VA accounted for the integration of 2D motion signals (e.g. plaids). However, IOC rule fails at predicting accurate direction in some cases (Champion, Hammett, & Thompson, 2007). Furthermore, adapting to gratings moving in the VA direction shifts the perceived motion toward the IOC direction and vice versa (Bowns & Alais, 2006). These results suggest that those rules should not be considered as independent alternatives and that results are often explained by a combination of the two. Finally, it is worth mentioning that Johnston and Scarfe (2013) have recently tried to disentangle the issue by introducing new integration rules that seem to compute perceived directions accurately.

Lorenceau et al. (2005) aimed instead at exploring the interactions between target and background signals. They presented observers with two vertical collinear line segments separated by a gap. According to four experimental conditions, a mask could either hide inner, outer, both or none segments line-ends. Segments simultaneously translated clockwise or counter-clockwise and observers had to report their perceived directions. Mask caused the segments to translate back and forth along a horizontal axis. Findings showed that, compared to a control condition, motion direction was harder to discriminate when the mask hid the outer ends, so that observers could only rely on inner ends. On the contrary, masking the inner-ends did not influence the performance. Interestingly, these results were satisfactorily replicated by a *cooperative-competitive* model involving long-range facilitation and surround suppression between collinear contours. The model is composed by two cell populations: one promoting integration and interacting via long-range horizontal connections with other cells having similar orientation preferences and aligned receptive fields (*I* cells, Figure 4c); the other responsible for segmentation and surround suppression (*S* cells, Figure 4c). *I* cells are activated by a collinear stimulus presented in their surround (*collinear facilitation*) and suppress the *S* cells sensitive to the same region of space (*end-stopping*). On the other hand, in the absence of a collinear stimulus, it is the *S* cells that are strongly activated and suppress the *I* cells. The decision related to the direction of the stimulus is mainly based on the activity of the S cells over time. Decision is made when the decision unit reaches a given threshold. According to the authors, this rather simple model follows the cortical distribution of long-range horizontal connections and predicts a release from the suppression of the responses of *S* cells sensitive to the inner line-ends when collinear facilitation is no longer provided, e.g. when a lateral offset is introduced between the two line segments, the relative difference in orientation of the two line segments is increased, the relative distance between the two line segments is increased. Although the architecture described above resembles that of area V1 (Hubel & Wiesel, 1968; Kapadia et al., 1995), the authors did not provide any specific neural substrate for their model, suggesting that motion integration could occur at different stages of the visual pathway.

Similarly to (Lorenceau et al., 2005), Vreven and Verghese (2002) also compared their data to predictions from a high-level cooperative system (Chang & Julesz, 1984; Grzywacz, Watamaniuk, & McKee, 1995; D. Williams, Phillips, & Sekuler, 1986; Yuille & Grzywacz, 1988). They also tested the spatiotemporal properties of the mechanism involved. They asked observers to detect an area of coherently moving dots, embedded in a distracters field. The target could differ from the background because of direction and speed of motion. The main manipulation consisted in varying the aspect ratio of the area containing the signal, without altering the signal-to-noise ratio. They found that detection was easier when the signal window was elongated in the direction of motion, for both speed- and direction- defined targets. Direction played again a prominent role in the integration of local signals, consistently with studies on moving contours (Bex et al., 2001, 2003; Ledgeway & Hess, 2002, 2006; Ledgeway et al., 2005). The model they proposed involved cooperativity among low-level motion detectors (Figure 4d). The cooperative mechanism described here is relatively simple: it assumes that low-level motion detectors are connected with other detectors in a spatiotemporal neighbourhood. Interaction, however, only occurs among detectors that are tuned to similar stimulus properties (e.g. direction). When a single detector is activated by its preferred stimulus, it in turn activates the linked detectors. Activation from neighbouring detectors with similar tuning, combined with activation from the stimulus, accounts for the enhanced performance for consistently moving signals. However, differently from similar models (Chang & Julesz, 1984; Grzywacz et al., 1995;

D. Williams et al., 1986; Yuille & Grzywacz, 1988) and (Lorenceau et al., 2005), the one presented by Vreven and Verghese (2002) only involves facilitation in the direction of motion (Snowden & Braddick, 1989; Yuille & Grzywacz, 1988), without the need for inhibitory connections. Also in this case, hints about the neural area mainly responsible for this type of integration are not provided. However, the lower time bound necessary to activate the mechanism is calculated to be about 200 ms and it can integrate signals over at least 3 deg (with other studies showing a spatial extent up to 10-20 deg, (Todd & Norman, 1995)). Thus the authors suggest that such a mechanism intervenes *after* the detection of each local signal.

The idea of a cooperative mechanism that integrates signals along the direction of motion is also in agreement with (Ledgeway & Hess, 2002).

To conclude, different classes of models have been presented in this short review. Each model tends to be stimulus- and task-dependent and it is particularly suited for the theoretical framework it is presented in. It is hard, then, to draw conclusions about a unique way to pool local motion signals. However, it should be underlined that they all give importance to *direction of motion*. I will come back on this issue in Chapters 4 and 5.

2.3. Visual Evoked Potentials (VEPs) in the study of Grouping and Motion Perception

The study presented in chapter 5 investigates the dynamics of motion-form interaction using VEPs. In this subsection of chapter 2 I will give an introduction to this technique accordingly. Moreover, mainly results from papers on VEPs and perceptual grouping and VEPs and motion will be summarized.

2.3.1. VEPs: a brief overview of the main components

The early components usually recorded from occipital sites span from ~60 ms (C1) to ~200 ms (N170) after stimulus onset and can have both positive (P1, P2) and negative (N1, N170) deflections (Steven J. Luck, 2005). The components, with their latency and polarity, are showed in Figure 5. Here follows a short description.

The C1 is the first main perceptual wave, it is largest at posterior midline electrode sites and seems to be generated from area V1 (Di Russo, Martinez, Sereno, Pitzalis, &

Hillyard, 2002; Foxe et al., 2008; Gomez Gonzalez, Clark, Fan, Luck, & Hillyard, 1994; Jeffreys & Axford, 1972). It generally arises 40–60 ms and peaks 80-100 ms after the stimulus onset (Steven J. Luck, 2005). Its polarity can vary in a way that it traces retinotopy of area V1: this component is positive for stimuli in the lower visual field and negative for stimuli in the upper visual field (Clark, Fan, & Hillyard, 1994; Jeffreys & Axford, 1972). Finally, C1 is highly sensitive to stimulus parameters, such as contrast and spatial frequency and geometrical properties within stimulus configuration, e.g. collinearity (Khoe, Freeman, Woldorff, & Mangun, 2004). Moreover, it seems to be strongly dependent on subjects' variability (Foxe & Simpson, 2002; Jeffreys & Axford, 1972; Molholm et al., 2002; Rademacher, Caviness, Steinmetz, & Galaburda, 1993). For this reason, C1 usually defined on an individual basis as the earliest deflection peaking before 100 ms post-stimulus onset (Kelly, Gomez-Ramirez, & Foxe, 2008).

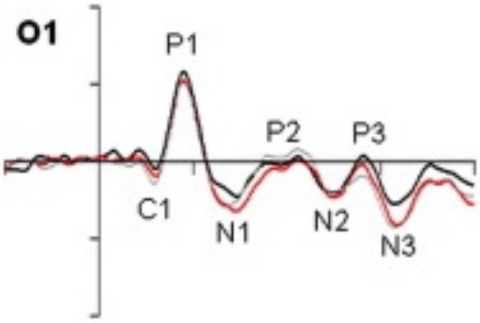


Figure 5. Representation of the main visual evoked components recorded from occipital electrode O1 in a ERPs paradigm (see text for details). The plot shows wave amplitudes (μV) as a function of time after stimulus onset (ms). The waves are labelled in a way that the letter indicates their polarity (P = positive; N= negative) and the number indicates their onset. C1 is the first component, peaking at 80-100 ms and can have either a positive or a negative polarity. P1 is the first positive component emerging 60-90 ms and peaking at 100-130 ms after stimulus onset and is largely thought to reflect physical features of the visual stimulus. The first negative component arising after P1 is N1 (peak at 100-150 ms), usually associated to texture segmentation. P2, N2, P3 and N3 are late components and are likely to reflect cognitive and endogenous processes, rather than responses to the physical properties of the stimuli. Figure published in (Casco, Campana, Han, & Guzzon, 2009).

The P1 wave shortly follows (and often overlaps with (Steven J. Luck, 2005)) C1. It is largest at lateral occipital electrode sites and typically onsets 60–90 ms and peaks between 100–130 ms post- stimulus. Note, however, that P1 latency varies substantially depending on stimulus contrast and on variations on other physical parameters of the stimulus, such as brightness (Johannes, Munte, Heinze, & Mangun, 1995). The P1 component is also sensitive to the direction of spatial attention ((Hillyard & Anllo-Vento, 1998); see reviews by (Herrmann & Knight, 2001; Hillyard, Vogel, & Luck, 1998)) and to the subject's state of arousal (Vogel & Luck, 2000), whereas other top-down variables were not generally found to influence it (Steven J. Luck, 2005). Using computational modelling and EEG-fMRI co-registration techniques, some authors (Di Russo et al., 2002) have attempted to localize the P1 wave source. They found that the early portion of the P1 wave arises from dorsal extrastriate cortex (in the middle occipital gyrus), whereas a later portion arises more ventrally from the fusiform gyrus. As pointed out by (Steven J. Luck, 2005), however, these findings need to be carefully considered given that several visual areas are activated within the first 100 ms after the onset of a visual stimulus, thus many of them might contribute to the voltages recorded in the C1 and P1 latency range.

The N1 wave follows the P1 wave and is formed by several visual subcomponents. The earliest subcomponent peaks 100– 150 ms at anterior electrode sites, and at least other two posterior N1 components typically peak 150–200 ms after the stimulus onset, one recorded from parietal cortex and another from lateral occipital cortex (LOC) (Steven J. Luck, 2005). Also studies on illusory contours have pointed out a N1 that might arise from the LOC (Foxe, Murray, & Javitt, 2005; Halgren, Mendola, Chong, & Dale, 2003; Knebel & Murray, 2012).

This component has widely been associated to texture segmentation: it is indeed elicited by the presence of a border between two adjacent textures that differ because of some features (e.g. local orientations, spatial frequency), despite no contour is explicitly present (Bach & Meigen, 1992; Caputo & Casco, 1999; Casco, Campana, Grieco, & Fuggetta, 2004; Casco et al., 2009; Casco, Grieco, Campana, Corvino, & Caputo, 2005). N1 components are finally influenced by higher-level processes, such as spatial attention (Hillyard et al., 1998; Mangun, 1995) and task demands (Hopf, Vogel, Woodman, Heinze, & Luck, 2002; Ritter, Simson, Vaughan, & Friedman, 1979; Vogel & Luck, 2000).

P2 is the component arising after the N1 wave. At posterior sites, this wave is often difficult to distinguish from the overlapping N1, N2, and P3 waves (Steven J. Luck, 2005). P2 recorded from these sites usually shows stimulus analysis that is no longer lined to its characteristics, rather this component indicates an endogenous processing (Machilsen, Novitskiy, Vancleef, & Wagemans, 2011; Straube & Fahle, 2010; Straube, Grimsen, & Fahle, 2010). A distinct P2 wave follows instead the N1 wave at anterior and central scalp sites. This anterior component is larger for stimuli containing target features (but only when the target is defined by fairly simple stimulus features (Steven J. Luck, 2005)), and this effect is enhanced when the targets are relatively infrequent (S. J. Luck & Hillyard, 1994).

To conclude this overview, the N170 is the last visual component and is a hallmark of face specialization: indeed, it is usually later and/or larger for inverted faces than for upright faces (Steven J. Luck, 2005). It is usually recorded at lateral occipital electrode sites, especially over the right hemisphere (Rossion et al., 1999). Even if inversion effects modulating N170 amplitude have been found with other stimuli (Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002; Schendan, Ganis, & Kutas, 1998), this wave is still considered an electrophysiological correlate of “face specificity” (Rossion & Gauthier, 2002).

2.3.2. VEPs: studies on contour integration

The temporal properties of the ERPs make them one of the most suitable tools to investigate dynamics of perceptual grouping. To date, several studies have investigated contour integration and ERPs. Most of them have consistently showed a modulation on the first negative component (N1, 70-150 ms after stimulus onset) correlated to the presence of a contour in a background. Moreover, they also agree in reporting absence of modulation on P1 amplitude (positive component that emerges at 60-130 ms after stimulus onset). Given that the presence of a contour only modulates late components, researchers suggest that perceptual grouping might occur before the N1 onset, i.e. before 150–160 ms. In other words, contour integration could be performed in two stages: in the first, local features are analysed (explaining lack of modulation on P1); in the second (corresponding to N1 latency) contour elements are grouped and differentiated from the background. Later components, such as P2 and P3, are usually linked to task demands and allocation of visual attention, rather than to physical attributes of the stimuli. This section will review some of these papers in detail.

A direct attempt to delineate the electrophysiological correlates of Gestalt grouping principles has been given by Casco et al. (2009). Observers were presented with triplets of Gabors embedded in a background of differently oriented elements. Behavioural and electrophysiological results in a global (judging the global orientation of the triplet) vs. local task (judging the local orientation of the central Gabor in the triplet) were compared. The local orientations of the target Gabors were varied, in order to favour grouping by good continuation and similarity (collinear and iso-oriented Gabors), similarity only (parallel and iso-oriented Gabors), or none (the Gabors were perpendicular one to the other). Whereas grouping generally led to higher accuracy, the two principles had two separate electrophysiological outcomes in the occipital sites. In the local task, their correlate was an increment in the amplitude of N1 (peaking at 150-250 ms for similarity and 75-250 ms for good continuation). In the global task, grouping by good continuation led to a positive deflection peaking at 40-179 ms after the stimulus onset. Conversely, a decrement in positivity at 275 ms was observed with similarity grouping. The authors concluded that grouping had two main effects: the first was to facilitate connections within the triplets, the second was to limit surround suppression from the background. The two effects had the same psychophysical results (i.e. enhancement of the target) and were disentangled by different electrophysiological correlates. Note that these results share similar principles to those proposed by (Lorenceau et al., 2005) (see above).

The electrophysiological correlates of contour detection seem to be context-dependent. Machilsen et al. (2011), indeed, recorded ERPs in two main conditions: the background was either classically made up of randomly oriented and positioned elements or composed by iso-oriented elements. The contour could be presented or not and observers were instructed to maintain fixation and to press a button in correspondence of catch trials. Machilsen and col.'s findings revealed that P1 amplitude was larger for random than for iso-oriented displays, regardless of the presence of the contour. Conversely, N1 amplitude was linked to the contour: it was larger and had shorter latency when the contour was embedded in the iso-oriented background. The authors also found a modulation on the P2 amplitude: it decreased in correspondence of contour stimuli, regardless of background configuration. They concluded that N1 might reflect both segmentation (as in previous literature, see above) and contour-integration components because it had a shorter latency when a contour was presented. P2 might instead reflect a cue-independent object representation, i.e., its amplitude is not *directly* influenced by stimulus features. In support of this characterization of P2 there is also a study by Straube and Fahle (2010); (Straube et al., 2010). Observers were asked to discriminate shapes that could differ from the background because of orientation cues, spatial frequency cues or both. Salience was varied varying the contribution of each of the two cues. In both studies, an increase in salience produced a decrease in P2 amplitude, showing that differences in the electrophysiological response reflect the *perceived* salience of the stimulus and not directly the underlying physical cue configuration.

Mathes, Trenner, and Fahle (2006): focus on electrophysiological correlate of contour integration and its interaction with task demands. They tested three Gaborised stimulus configurations (CƆ, CC and ƆC) and manipulated the local orientations of the Gabors forming the targets to be either aligned or misaligned (a slight orientation jitter was applied). Also a control condition in which the background only was presented was used. To control for task demands, a masking paradigm was chosen. They designed two experiments that differed by the presentation order of stimulus and mask display: in the first experiment, the contour display was preceded by a masking window; in the second one, the mask followed the stimulus display. In both experiments, they found that control conditions did not modulate the N1 component, whereas higher N1 amplitude was found with contour displays. This effect started about 150 ms after stimulus onset on the posterior sites, but was delayed for more difficult detection conditions. Moreover, the presence of the target elicited a P3, which was larger for aligned compared to jittered contours. The main difference between the two experiments consisted in the observation of a frontal selection positivity (FSP) elicited by presenting the stimulus after a random display (experiment 1). Consistently with other studies (Casco et al., 2009; Machilsen et al., 2011), they concluded that the early negative shift might reflect similar processes in contour integration and texture segmentation. With decreasing salience (e.g., as orientation jitter increases or with particular stimulus configurations), contour integration might need the allocation of selective attention, thus explaining the increasing processing time found in this study.

The dynamics and loci of contour processing were also tested by Tanskanen, Saarinen, Parkkonen, and Hari (2008) using MEG. The authors were interested in assessing i) when the difference between displays containing background elements only and background and contour emerges; ii) what are the loci of early-contour processing; iii) and what effects local orientations have on cortical responses. They presented observers with two concentric contours, formed by either tangential or radial elements. The contours could either be a complete circumference or part of it, occupying a single sector. They found that displays containing a contour started differentiating at 130 ms from the stimulus onset. Source analysis revealed that responses were located at high areas in the visual hierarchy (i.e. V6). Moreover, cortical response was stronger for complete and tangential contours. They concluded that extrastriate areas are involved in global shape integration. Results are consistent with another study (Pei, Pettet, Vildavski, & Norcia, 2005) showing that response to shapes started at 130 ms after stimulus onset and was accompanied by larger N1 amplitude.

Another approach comes from studies on crowding. According to a growing body of evidence, perceptual grouping and crowding are two sides of the same coin (May & Hess, 2007b; Pelli et al., 2004). Nonetheless, only few studies on ERPs and crowding have been published. Among them, the study by Chicherov, Plomp, and Herzog (2011) measured the visual components evoked by a Vernier task in three conditions: the Vernier target could be surrounded by shorter, equal length and longer flankers on both sides. The rationale came from a previous study showing a release of crowding when flankers were grouped together, i.e. when their length was different from target's one (Malania, Herzog, & Westheimer, 2007). They found modulation of two main components: P1 amplitude monotonically increased with target size, indicating extraction of basic perceptual features. N1 amplitude was smaller with shorter flankers and increased for same length and longer flankers. They concluded that grouping happened somewhere after P1 and before N1 because the flanker length effect was only visible on N1 amplitude. These findings are consistent with previous results showing the crucial role of the N1 component in figure-ground segregation tasks (see above).

2.3.3. VEPs: studies on motion perception

Visual Evoked Potentials are a powerful and relatively undemanding way to study the visual system. It is not surprising then that this method has been used also in the study of motion perception (see (Heinrich, 2007) for a specific review). Several investigations were dedicated to clinical aspects, given that motion perception deficits are often linked to pathologies (Kubova & Kuba, 1995; Kubova, Kuba, Peregrin, & Novakova, 1996; Norcia et al., 1991). However, other works have focused on the electrophysiological correlates of motion perception itself. Here it follows a brief overview of some of these studies.

To investigate motion, researchers have taken advantage of two known properties of the mechanisms underlying its processing, i.e. direction-selectivity and adaptation, which also have segregated VEPs correlates.

Initial studies focused on stimuli that started to move after being stationary. Results from these studies revealed that *motion-onset* VEPs are characterised by an occipito-temporal negativity that peaks between 150-200 ms after the motion onset (N2). An interesting peculiarity of the N2 is that is lateralized, with higher amplitude recorded in the left hemisphere in 60% of subjects (Kubova, Kuba, Hubacek, & Vit, 1990). Motion-offset too seems to elicit a negative deflections, although smaller than N2 (Kuba & Kubova, 1992). Further, stimuli at low-contrast (Bach & Ullrich, 1997; Kubova, Kuba, Spekreijse, & Blakemore, 1995; Muller & Gopfert, 1988) and direction-dependent adaptation responses (Heinrich & Bach, 2003; Hoffmann, Unsold, & Bach, 2001; Muller, Gopfert, Leineweber, & Greenlee, 2004; Wist, Gross, & Niedeggen, 1994) provoke modulation of this component.

Bach and Ullrich (1997) compared the contrast dependency of VEPs to motion onset at an occipital and lateral recording site. Their stimuli were sine-wave grating ideally set to elicit a motion response, because they drift at intermediate speed (4.9 deg/sec) and had a low spatial frequency (0.9 c/deg). Contrast ranged from 0.4 to 98%. Observers performed an unrelated task, while fixating. Results showed a negative component at around 180 msec (N2), recorded from lateral sites, enhanced in motion-onset stimulation, exhibiting a low contrast threshold (2%). Such negativity saturated for contrast values above 4%, demonstrating higher sensitivity for lower contrast.

As anticipated, N2 is tuned to direction, with different directions eliciting different modulations: complex motion (e.g. radial motion) produces higher amplitude (Kremlacek, Kuba, Kubova, & Chlubnova, 2004), especially with radial expansion (Holliday & Meese, 2005). Translational motion also seems to induce an effect: using an orthogonally drifting grating, Maffei and Campbell (1970) showed lower amplitudes for oblique compared to cardinal motion directions. These effects might be related to different patterns of activation of parietal and extrastriate areas caused by different moving stimuli (Delon-Martin et al., 2006). Direction tuning is very broad, with the bandwidth of an estimated motion detector spanning a range as wide as 62° (recorded at occipito-temporal electrodes) (Maurer, Heinrich, & Bach, 2004).

N2 is also speed tuned, with high speeds leading to decrement in latency and increment in amplitude of the component (Heinrich, 2007). Studies combining psychophysical methods, such as adaptation, and VEPs recording give support to the idea that there are different channels in the visual system tuned to different speed bandwidths (Heinrich, van der Smagt, Bach, & Hoffmann, 2004). It is not clear yet, however, whether this tuning is independent of the spatial properties (e.g. spatial frequency) of the stimuli (Heinrich, 2007).

Finally, contributes of cognitive aspects (e.g. attention and task demands) have been also investigated. Results on the role of attention are yet not clear. On one hand, when observers are asked to attend static instead of dynamic lines in a transparent motion display, the related N2 amplitude is smaller (Torriente, Valdes-Sosa, Ramirez, & Bobes, 1999). On the other hand, N2 amplitude increases when infrequent (compared to frequent) motion direction is presented, even if attention is consciously directed to a neutral stimulus (Pazo-Alvarez, Amenedo, & Cadaveira, 2004). Task-demands also influence N2: detection tasks yield shorter latencies at offset, onset and peaks than identification tasks (Fort, Besle, Giard, & Pernier, 2005).

Source-analysis localised N2 origin in area MT (Amano, Kuriki, & Takeda, 2005; Kaneoke, Watanabe, & Kakigi, 2005; Nakamura et al., 2003; Probst, Plendl, Paulus, Wist, & Scherg, 1993), although also contributes from areas V3/V3A might be present (Bundo et al., 2000; Schellart, Trindade, Reits, Verbunt, & Spekreijse, 2004). The transition from one complex to the other might result in 20 ms delays on the component's latency (Bundo et al., 2000; Schellart et al., 2004). The idea that N2 is generated by area MT is also supported by studies on random dots. Coherence paradigms indeed selectively tap this area. Since N2 is modulated in latency by increase in coherence in this task, it is considered to mirror activation in area MT (Amano et al., 2006; Aspell, Tanskanen, & Hurlbert, 2005).

Nonetheless, using VEPs, ffytche, Guy, and Zeki (1995) found evidence that MT might be (at least partially) a stage of a motion processing path that goes through or passes by V1. The authors investigated the spatiotemporal characteristics of very early (35–105 ms) cortical responses to motion. They found that slow motion activated first V1 and subsequently MT, while fast motion directly activated MT, thus suggesting the presence of two pathways to MT.

To conclude, VEPs in visual motion tasks have revealed a component, N2, which is usually associated to motion onset. Other components (such as P1) might be related to motion, but results are in contradiction (Heinrich, 2007). Moreover, it has been shown that N2 might originate in area MT, consistently with physiological studies (Anderson, Holliday, Singh, & Harding, 1996). Accordingly to psychophysical reports, this component is influenced by direction, speed and contrast of the eliciting stimulus and is susceptible to adaptation.

Chapter 3. The integration of straight contours (snakes and ladders): The role of spatial arrangement, spatial frequency and spatial phase.

1. Introduction and aim of the study

To recognize a contour made up of disconnected elements, these elements have to be correctly linked together and segmented from those not belonging to the contour. Since linked elements are locked in their orientation (local) to the curvature of the whole contour path (global), they respect the Gestalt law of "good continuation." This law states that discrete contour elements, positioned and oriented along a smooth path, are readily grouped together into a salient contour by a process that establishes a relationship between local and global orientation (Wagemans, 2012; Max Wertheimer, 1938). Field et al. (1993) formalized the "good continuation" law in the "association field" model, according to which binding follows a joint constraint of position and orientation.

The association field model does not include specific postulations to account for a combination of good continuation with another Gestalt rule (Max Wertheimer, 1938) of grouping, that of "similarity." This rule is optimally expressed within straight contours made up of iso-oriented (and thus similar) elements with either a collinear (snakes) or a parallel arrangement (ladders). Similarity may indeed play a role, together with good continuation, in the integration of straight contours and could account for three well-documented findings. The first is that straight contours (formed by iso-oriented patches) are more salient than curvilinear ones (Field et al., 1993; R. F. Hess & Dakin, 1997, 1999; Ledgeway et al., 2005). The second is that, although ladders do not respect the joint constraint of position and orientation, they are still well detected (Bex et al., 2001; Dakin & Baruch, 2009; Hansen & Hess, 2006; Ledgeway & Hess, 2006; Ledgeway et al., 2005), and this might be due to the fact that elements in a *ladder* share the same local orientation. The third is that snakes are generally more salient than ladders (Bex et al., 2001; Dakin & Baruch, 2009; Field et al., 1993; Ledgeway & Hess, 2006; Ledgeway et al., 2005; May & Hess, 2007a, 2007b, 2008), being snake contours defined by both iso-orientated and collinear elements (Casco et al., 2009; Kapadia et al., 1995; W. Li & Gilbert, 2002).

In the present study I asked two questions. The first is whether straight contours are more salient than predicted by the classical association field because they involve grouping by similarity (e.g. iso-orientation), in addition to grouping by good continuation (mainly driven by collinearity). I addressed this question by manipulating the spatial arrangement of the elements in the background (Polat & Bonneh, 2000), thus introducing a contextual influence on contour binding. Indeed, we positioned the elements of the background on a grid. This way they are perceived as grouped into stripes regardless of their orientation (Bozzi, 1969, 1989; Max Wertheimer, 1938). This role of "pure" good continuation on grouping has been described by Gestalt psychologists (Bozzi, 1969, 1989; Max Wertheimer, 1938) who have thought it the spatial analogy of common fate (Bruce, 2003) and its neurophysiological correlate could rely on the connections between pairs of neurons of all relative orientation preference (Das & Gilbert, 1999). I expected grouping of randomly oriented background elements into stripes to have a strongly suppressive contextual effect on detection of the target "stripe," i.e. the straight target contour. Note that this effect should not be confused with recently discovered local contextual effects in contour integration (Dakin & Baruch, 2009; May & Hess, 2007b; Robol et al., 2012; Schumacher et al., 2011).

Two background conditions were set: (1) background elements were randomly positioned, with the contour distinguished because of both good continuation and similarity of its elements, (2) background elements were positioned orderly in a grid, with detection mainly based on similarity. I predicted that if only good continuation underlies detection (as with curvilinear contours), the contextual effect of the orderly-positioned background should prevent detection. On the other hand, if both grouping rules (good continuation and similarity) were used by the association field, the orderly-positioned background should reduce but not prevent detection.

The second question is whether integration of straight contours occurs within a receptive field or across receptive fields. It has indeed been suggested that facilitatory interactions *within* one receptive field may account for detection of iso-orientation in ladders (N. Graham & Wolfson, 2004) and snakes (Levi & Waugh, 1996; Yu & Levi, 1997). Other authors instead suggest that contextual interactions and intrinsic horizontal connections *between* neurons in the primary visual cortex might underpin contour integration (W. Li & Gilbert, 2002). I sought to distinguish between these two explanations thanks to the comparison of the contextual effects produced by the two types of background. Indeed, contextual influence of the orderly-positioned background is not easily accounted for by the influential models of straight contours detection which assume that integration occurs within one high-level filter (R. F. Hess & Dakin, 1997, 1999; May & Hess, 2007b, 2008). These mechanisms may respond to feature contrast (N. V. Graham, 2011) in the centre and the surround of their receptive field (mainly to orientation but also to contrast polarity and spatial frequency), regardless of the spatial arrangement of these features. A suppressive contextual effect exerted by the regular background would instead be compatible with integration based on intrinsic long-range horizontal connections. These connections are formed in V1 by axons of pyramidal cells and linking neurons with non-overlapping receptive fields and the same orientation preference (Ts'o & Gilbert, 1988; Ts'o, Gilbert, & Wiesel, 1986). These connections are mostly excitatory for both collinear (Kapadia et al., 1995) and parallel receptive fields (Fitzpatrick, 1996; Mitchison & Crick, 1982; Rockland & Lund, 1982, 1983), but inhibitory interactions between iso-oriented parallel elements are also possible (Z. Li, 1998; McGuire, Gilbert, Rivlin, & Wiesel, 1991; Polat, 1999).

In addition, to explore further the role of lateral interaction in straight contour integration, the phase of the elements in the target was varied. Other studies have varied inter-element distance and element density (W. Li & Gilbert, 2002; Polat, 1999; Polat & Bonneh, 2000) and phase (Bell, Gheorghiu, Hess, & Kingdom, 2011; C. B. Williams & Hess, 1998) to this end. In particular, binding of disconnected elements to curvilinear contours is not strongly affected by phase variation (Field et al., 2000), supporting the view that this operation involves high-level visual processes. Facilitation by collinearity in contrast detection tasks also shows relative phase insensitivity in some conditions (Chen & Tyler, 1999; Wehrhahn & Dresp, 1998; Yu & Levi, 1997). Other studies, however, have shown a strong effect of phase in detection of aligned iso-oriented micropatterns, leaving open the possibility that although phase does not play a role in integration of smoothly curved paths, it could affect integration of straight paths (C. B. Williams & Hess, 1998). My hypothesis is that, if manipulating phase (and spatial organization in the background) produced an effect on snake and ladder detection, this would suggest that integration involves intracortical lateral interactions at early level filters (Bonneh & Sagi, 1998). Phase insensitivity and absence of contextual effects would instead support integration within a single receptive field of second-stage filters after rectification of first-stage filters (Chen & Tyler, 1999; Usher, Bonneh, Sagi, & Herrmann, 1999).

2. Experiment 1

Experiment 1 compares detection of straight contours embedded in a surround of either randomly-positioned elements or elements positioned on an orderly grid (orderly-positioned).

2.1. Methods

2.1.1. Subjects

The author and five naïve participants served as observers in Experiment 1. Observers sat in a dark room and were placed 57 cm from the screen. Viewing was binocular. They were given initial training to familiarize themselves with the stimuli and the task. All subjects had normal or corrected-to-normal visual acuity. Subjects participated voluntarily with no compensation and gave their informed consent prior to their inclusion in the experiment.

2.1.2. Apparatus and stimuli

Stimuli were displayed on a 19-in. CTX CRT Trinitron monitor with a refresh rate of 60 Hz and generated with Matlab Psychtoolbox (Brainard, 1997; Pelli, 1997). The screen resolution was 1280 × 1024 pixels. Each pixel subtended ~1.9 arcmin. The mean luminance was 32 cd/m^2. Luminance was measured with a Minolta LS-100 photometer. Stimuli were presented within a square window (10 × 10°) placed at the centre of the screen. Each stimulus was generated anew immediately prior to its presentation and consisted of a dense spatial array of 100 Gabor patches. Each Gabor was defined as follows:

$$G(x, y) = m * sin(2\pi f x x \pi + \varphi) * exp(-(x2 + y2)/\sigma 2)$$

where m is the modulation depth of the sinusoidal carrier, *fx* is the spatial frequency (SF), ϕ is the phase of the carrier (0°). The Gaussian is expressed by the exponential part of Eq. (1): x and y are, respectively, the horizontal and the vertical distance from the peak of the Gaussian, whereas σ is the standard deviation (0.21°). The contrast was fixed at 0.5 (Michelson contrast).

Gabor patches were placed within a raster of 10 × 10 cells (each cell was 1 × 1°) (Field et al., 1993). Each Gabor patch occupied a single cell. The target path was made up of five elements with orientation either parallel (snake, Fig. 1, Panel A) or orthogonal to the global orientation of the path (ladder, Fig. 1, Panel B). In addition, the path could be either horizontal or vertical. Thus, snake paths were defined by elements with the same element and path orientation, whereas ladders were defined by elements with the same orientation but orthogonal to the path's orientation. The background elements had a random spatial orientation (ranging from zero to 360°). In addition, they could be either orderly positioned on a grid (orderly-positioned condition, Fig. 1, bottom row in Panel A and Panel B) or positioned randomly, by spatially jittering their ordered position in a random direction (ranging from zero to 360°) with respect to the centre of the cell by 0.3° (randomly-positioned condition, Fig. 1, top row in Panel A and Panel B).

In the latter condition, Gabors could slightly overlap, one occluding the other without summing their luminance profile; in this way we prevented changes in luminance to create cues.

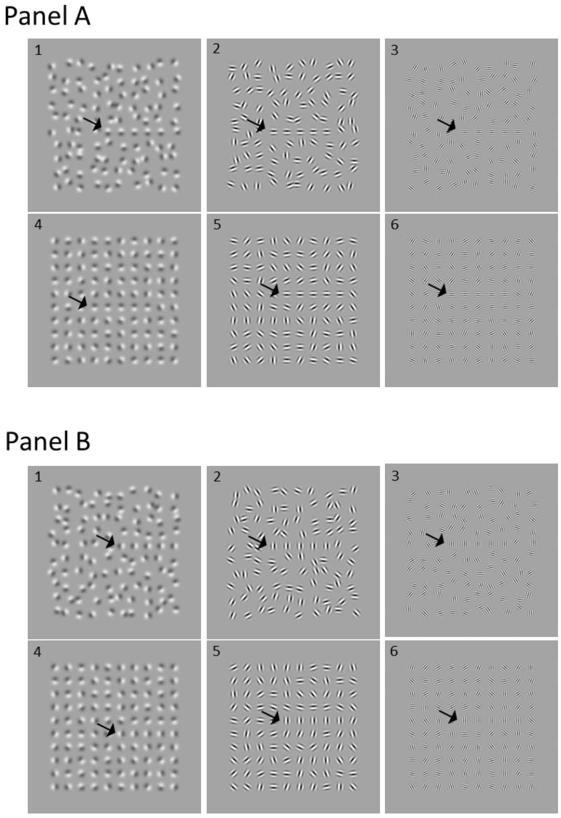
In the case of the randomly-positioned condition, mean luminance and contrast of the Gabor patches were the same even when they slightly overlapped.

Distance between elements was 1° (centre-to-centre distance) and on average it was maintained approximately constant in the randomly-positioned condition as well. Also, since the background elements were randomly oriented, the context did not create specific spatial cues for either ladder or snake paths.

2.1.3. Procedure

A two-interval forced-choice (2IFC) procedure was used. In each trial observers were required to choose which of two temporal intervals (followed by 100 ms mask and an Inter-Stimulus-Interval [ISI]: 100 ms) contained the target path.

For example, the target path, either a snake or a ladder, was presented in the first interval embedded in the background whereas, in the second interval, only the background was displayed, or vice versa. The interval in which the target path appeared was randomized across trials. During the ISI, the screen was set to the mean luminance (32 cd/m^2).



Panel A

Panel B

Figure 1. The Figure shows the stimuli used in Experiment 1. The elements forming the contour (indicated with the arrows) were orderly positioned and had a constant phase and the same orientation (either parallel to the path: snake [Panel A] or orthogonal to the path: ladder [Panel B]). The elements in the background had random position and orientation (randomly-positioned background condition [Panel A and B, from 1 to 3]). Panels A and B from 4 to 6 show the stimuli used in the orderly-positioned background condition. The three spatial frequencies tested (1 cycles/°, 3 cycles/° and 6 cycles/°) are shown in columns. In the experiment, the vertical contour was also presented. For demonstrative purposes, the figures show stimuli with an exaggerated contrast.

In order to reduce the spatial uncertainty of the target path, it was displayed and constrained to lie within a central squared region of the stimulus window (7.5 × 7.5°). The orientation of the whole path (i.e. vertical vs. horizontal), the stimulus type (snake vs. ladder) and the duration of each interval (from 50 to 300 ms: 50 ms step) were

varied within each block. The type of background (i.e. orderly- vs. randomly-positioned) and the SF (1, 3 and 6 cycles/°) were varied across blocks.
For each interval the stimulus was immediately followed by a mask made up of a raster of Gabor patches with random orientation, random phase value and either aligned (in the orderly-positioned condition) or spatially jittered (in the randomly-positioned condition). There were a total of 52 conditions: 2 background conditions × 2 types of target path (i.e. snake vs. ladder) × 6 stimulus exposures (from 50 to 300 ms: 50 ms step) × 3 spatial frequencies (i.e. 1 cycles/° vs. 3 cycles/° vs. 6 cycles/°). Observers performed 32 repetitions per condition.

Individual performances of three subjects in a pilot experiment conducted with Gabor patches with a spatial frequency of 1 cycles/° were used to estimate the appropriate range of exposures that would avoid ceiling and floor effect.

A logit function (Berkson, 1953; Chatterjee & Price, 1977) was fitted to the data in order to estimate the thresholds, defined as the exposure level at which observers detected the target path with a probability of 0.75. The two psychometric functions, describing accuracy as a function of exposure (ranging from 20 ms to 640 ms) separately for snakes (continuous line, Fig. 2) and ladders (dashed line, Fig. 2), showed a proportional increase in accuracy from 50 ms to 300 ms. Thresholds were higher for the three subjects in the ladder condition (171.1 ms for subject CV, 283.1 ms for subject NL and 311.1 ms for subject CO) than in the snake condition (89.36 for subject CV, 158.7 ms for subject NL and 221.7 ms for subject CO). Moreover, the two psychometric functions had similar slopes (0.008 with snakes and 0.005 with ladders for observer CV, 0.006 and 0.004 for NL, 0.003 and 0.001 for CO), indicating that the difference between the two contours was constant over a range of exposures. Thus, for all the subjects, performance increased proportionally with duration. These exposure levels were taken as the range to be used in the following experiments.

2.1.4. Results and discussion

Results from Experiment 1 are illustrated in Fig. 3. Overall, the results show that detection of snakes is high for all the SFs in the condition where background elements are randomly positioned and improves gradually with increasing SF in the regularlypositioned background. Thus, the detection of snakes is more accurate with the randomly-positioned background only at low SF. For ladders, instead, detection is very impaired at low SF in both the background conditions, and improves more with increasing SF in the orderly-positioned condition. At the highest SF, ladder detection is better with the randomly-positioned background. Since detection increases proportionally with the spatial frequency employed (Fig. 3), only the results with the highest and the lowest spatial frequency were analysed for better assessment of this effect.

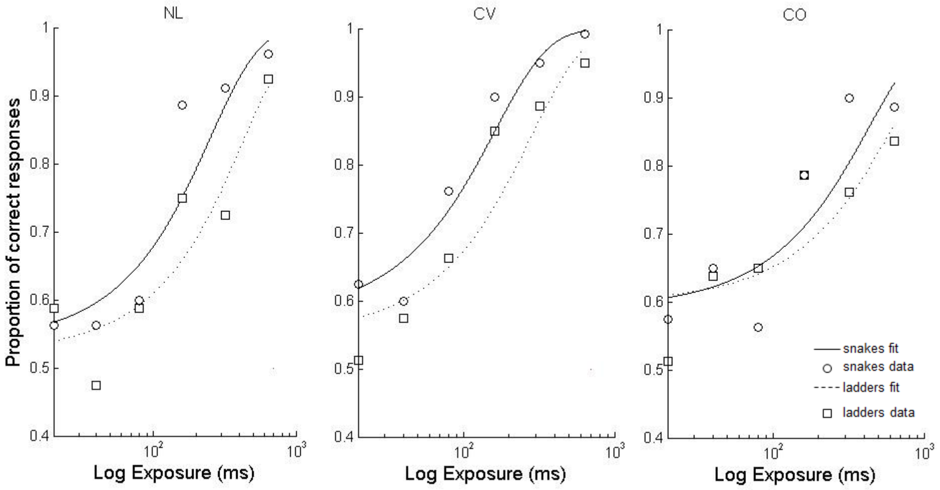


Figure 2. Individual psychometric functions of the three subjects participating in the preliminary study. Circles refer to snake detection and squares refer to ladder detection. Stimuli were randomly presented at 20 ms, 40 ms, 80 ms, 160 ms, 320 ms and 640 ms. As expected, performance increased as a function of the exposure in both the conditions.

The complex dissociation of the effect of the background on the detection of snakes and ladders was pointed out by two-way repeated-measures ANOVAs with Background Type (orderly- vs. randomly-positioned) and Exposure (100, 150, 200, 250, chosen to reduce ceiling and floor effects) as factors. ANOVAs were run separately for the two contour types and two spatial frequencies (1 and 6 cycles/°).

For snakes, besides the significant effect of exposure at high SF ($F(3,5) = 18.431$, $p = 0.004$), the predicted influence of the background was revealed by a significant effect of this main factor ($F(1,5) = 9.532$ and $p = 0.027$) at low SF and by a significant Background × Exposure interaction ($F(3,5) = 4.428$, $p = 0.038$) at high SF. In the latter the randomly-positioned background only improves performance at an exposure duration of 250 ms (Bonferroni corrected t-test, p = 0.025).

Conversely, for ladders, a main effect of Background type was only present at 6 cycles/° (F(1,5) = 16.286, p = 0.010). The effect of exposure was also significant at both low (F(3,15) = 8.288, p = 0.008) and high (F(3,15) = 12.950, p = 0.001) SF. None of the other main effects or interactions were significant.

Considered together, the results of Experiment 1 confirmed that both snakes and ladders are very salient even at short exposures (Ledgeway et al., 2005; May & Hess, 2007a).

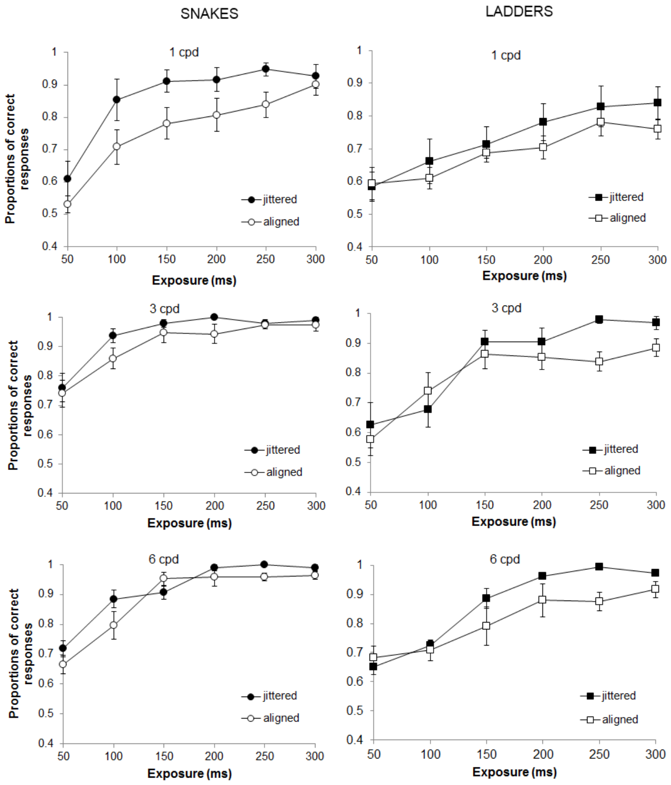


Figure 3. Results of Experiment 1. Proportions of correct responses were plotted as a function of the exposures. The solid symbol refers to the randomly-positioned background condition and the empty symbol to the orderly-positioned background condition. Results for snakes and ladders are shown in the left and right column, respectively. The three spatial frequencies used were plotted separately (1 cycles/° in the top row, 3 cycles/° in the middle row, 6 cycles/° in the bottom row). SEs are plotted for each data point.

The new finding is a different effect of the grid-positioned background on snake and ladder detection specific to the spatial frequency tested: snake contours are impaired by the orderly grid-positioned background at a low spatial frequency, whereas ladder contours are more salient with the randomly-positioned background, at a high spatial frequency.

3. Experiment 2

Whereas binding of curvilinear contours is little affected by phase manipulation (Field et al., 2000), some studies have reported a strong effect of phase on collinear facilitation (C. B. Williams & Hess, 1998). If phase played a role, this would suggest that detection occurs at the level of first-order filters before a full-wave rectification of the output (Chen & Tyler, 1999) and that lateral interactions between them may occur (Bonneh & Sagi, 1998).

The role of phase was investigated with the two types of background, since phase randomization of elements in the target path may interact differently with the contextual influences of randomly-positioned and grid-positioned background elements.

3.1. Observers

The same observers who took part in Experiment 1 participated in Experiment 2.

3.2. Apparatus and stimuli

The general method used in this experiment was identical to that employed in Experiment 1, with the exception that the elements in the target path could assume, randomly, two phase values. To produce Gabors with randomly alternated phase, the phase value of each element in the path was randomly assigned between 0° or 180°. Gabors in the path had the same orientation, as in Experiment 1, either forming a snake (Fig. 4, Panel A) or a ladder contour (Fig. 4, Panel B). The same two background conditions of Experiment 1 (i.e. randomly- vs. orderly-positioned) were used (Fig. 4).

3.3. Procedure

The same procedure as in Experiment 1 was used. We tested the SFs of 1 and 6 cycles/°.

3.4. Results and discussion

Results from Experiment 2 are shown in Fig. 5. The main finding is that phase randomization selectively impairs snake detection at low spatial frequency. Conversely, ladder detection shows no effect at low spatial frequency and it slightly but significantly improved at high spatial frequency when phase is randomized. Moreover, this effect was present for all the exposures tested.

We confirmed these observations by means of three-way repeated measure ANOVAs with Phase (constant, Experiment 1 vs. random, Experiment 2), Background Type (randomly- vs. orderly-positioned), and Exposure (100, 150, 200, 250 ms) as factors. Analyses were run separately for snakes and ladders and for 1 cycles/° and 6 cycles/°.

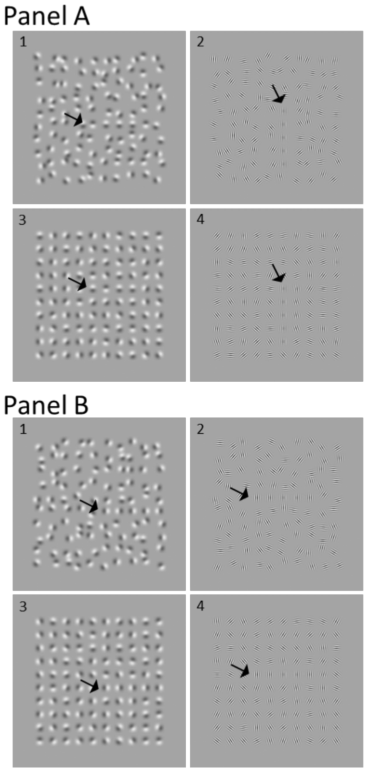


Figure 4. Examples of the stimuli used in Experiment 2 (random phase condition). The elements forming the contour (indicated by arrows) had regular positions and they all had random phases and the same orientation. Both snakes (Panel A) and ladders (Panel B) were presented. In the random-positioned background condition, the elements in the background had random position and orientation (Panels A and B, Fig. 1 and Fig. 2). In the orderly-positioned background condition (Panels A and B, Fig. 3 and Fig. 4) the elements in the background had random orientation and regular position. The two spatial frequencies tested (1 cycles/° and 6 cycles/°) are shown in columns. The contours could be either vertical or horizontal. For demonstrative purposes, the figures show stimuli with an exaggerated contrast.

For snakes, we found a significant effect of Phase (F(1,5) = 56.131, p = 0.001) only at low SF. As in Experiment 1, we found an effect of Background (F(1,5) = 35.559, p = 0.002) at low SF and a significant interaction Exposure × Background Type (F(3,15) = 4.358, p = 0.039) at high SF, showing that performance increased with the random background only for the highest duration of the stimulus (p = 0.030, Bonferroni corrected t-tests). Not surprisingly, we found a significant effect of Exposure, both at low (F(3,15) = 10.421, p = 0.001) and high (F(3,15) = 23.445, p = 0.002) SFs.

ANOVAs on ladders also revealed a different pattern of results at low and high spatial frequency. Indeed, we only found a significant effect of Exposure when testing 1 cycles/° (F(3,15) = 8.962, p = 0.008), whereas the three principal effects were found at high spatial frequency (Phase, F(1,5) = 10.970, p = 0.021; Background, F(1,5) = 19.694, p = 0.007; Exposure, F(3,15) = 22.295, p = 0.0001), showing an increase in detection with, interestingly, randomized phase, random background and increase in exposure duration. None of the other main effects or interactions were significant.

Taken together, the results revealed that phase manipulation selectively reduces snake detection, but only when the spatial frequency tested is low (i.e. 1 cycles/°). It also improves ladder detection at high spatial frequency for both the backgrounds tested.

Note that previous studies on contour binding failed to show severe impairment in detection as a consequence of phase randomization (Field et al., 1993, 2000; R. F. Hess & Dakin, 1999).

4. General discussion

To summarize, the results showed that detection of snakes defined by Gabors with low spatial frequency carriers was hampered by orderly-placed background elements in a grid. Overall, ladders were less salient than snakes and, when defined by Gabors with high spatial frequency carriers, their detection was improved by random positioning of the background. Moreover, regardless of the spatial organization of the background, snake detectability increased solely because of phase constancy, specifically at low spatial frequency, whereas ladders were slightly but consistently improved by phase randomization.

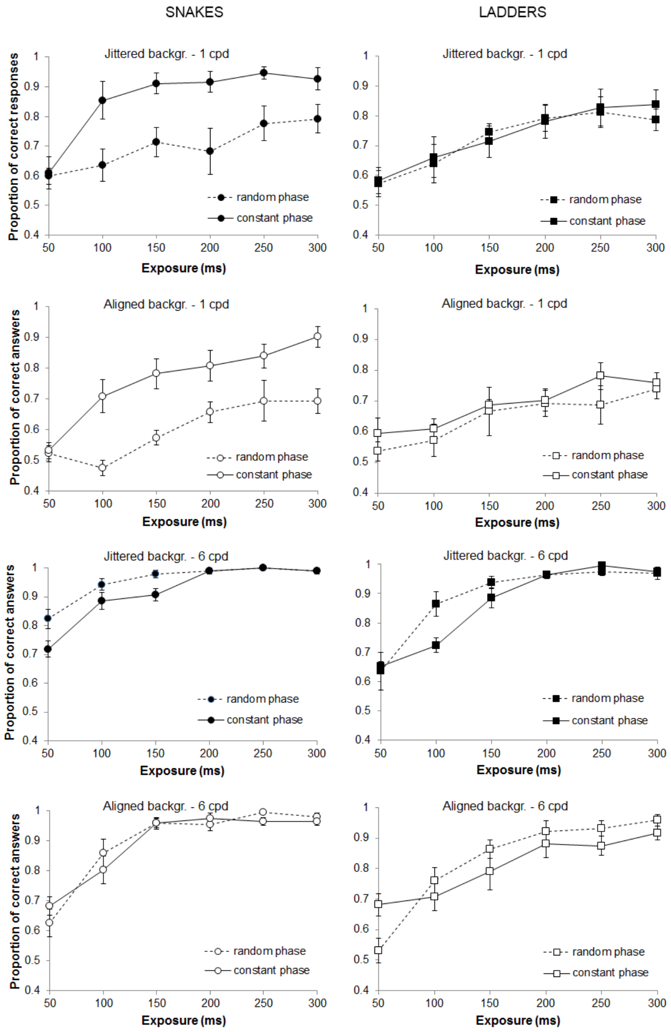


Figure 5. Results of Experiment 2. Proportions of correct responses are plotted as a function of the exposures. SEs are plotted for each data point. In different rows, data for the comparison between the constant-phase condition (continuous line) and the random-phase condition (dashed line) are shown. The solid symbol refers to the randomly-positioned background condition and the empty symbol to the orderly-positioned background condition. Results for snakes and ladders are shown in the left and right column, respectively. The two spatial frequencies used were plotted separately (1 cycles/° in the top, 6 cycles/° in the bottom).

We discuss the combined effect of phase randomization and spatial arrangement of background separately for snakes and ladders, with the aim of determining whether the integration of iso-oriented elements into the two types of contours can be accounted for by lateral interactions between local elements.

Note that we assume that interference may occur between mechanisms with the same response properties operating at the same stage of processing, either at the first or at a higher level.

4.1. Snakes

Detection of snakes could in principle be accounted for by integration within a single receptive field, either at a first or at a second stage. Our findings, however, do not support this view. Integration within a single second-stage receptive field is, on the one hand, not compatible with the phase effect, when this filter is fed by the full-way rectified output of first-order filters. On the other hand, the output of a second-stage filter fed by the output of half-way rectified first-stage filters would be affected by phase perturbation (Motoyoshi & Kingdom, 2007). This filter, however, would be insensitive both to iso-orientation, because it integrates across different carrier orientations (Motoyoshi & Kingdom, 2007), and to the spatial arrangement of the background where the random orientation and phase produce zero output in this mechanism.

At a first stage, low spatial frequency Gabors forming the straight contour optimally stimulate a detector of luminance gradient with odd receptive field, provided that the dark region of the Gabor patch falls onto the inhibitory portion of the receptive field and the light one onto the excitatory portion. Obviously such a mechanism does not respond to a contour made up of phase randomized Gabors, and indeed we found that at low stimulus duration where scrutiny and eye movements are not allowed detection probability is far below 0.75 (the threshold value). In addition, such a first-stage filter returns zero output in the background where elements are randomly oriented, a fact which is incompatible with the impairment owing to the spatial arrangement of the surround elements orderly positioned on a grid.

An interesting model which could account for the failure in detecting snakes, when the phase is randomized, is that which assumes the activation of first-stage luminance filters that use position as the feature involved in contour processing (Prins, Kingdom, & Hayes, 2007). This mechanism assesses the shape of the straight contour based on the local positions of Gabors with the same phase. Researchers (Dakin & Hess, 1999; Hansen & Hess, 2006; Phillips & Todd, 2010) have pointed out the importance of phase alignment, especially in terms of texture discrimination. In the randomized phase condition this mechanism may produce local tilts, since it is possible that the phase of the Gabor carrier affects its perceived location. Randomizing the phase of the Gabor would then introduce positional noise and make it more difficult to perceive the collinearity of elements in the straight snake. A mechanism using position as a feature involved in contour processing is not sensitive to local orientation (Prins et al., 2007) and this could explain the detrimental effect on detection when, in an ordered-grid background, the only cue is orientation gradient. The manner in which this mechanism assesses the shape of the straight contour based on the local position of Gabors with the same phase has not been determined yet. It cannot be ruled out that this occurs through lateral interactions between first-stage mechanisms with the same phase (Roncato & Casco, 2006, 2009) that have been shown to be extremely good at detecting relative position (R. J. Watt & Andrews, 1982).

4.2. Ladders

Detection of ladders fails at low spatial frequencies regardless of phase constancy and background spatial arrangement. Ladders are detected rather when defined by high spatial frequency Gabors, particularly when the background elements are randomly positioned and the phase of elements in the contour is random. The inability to detect low spatial frequency ladders is particularly evident at short stimulus duration (detection probability is below 0.75, the threshold value). The improvement by phase randomization excludes the possibility that detection can be accounted for by an integration mechanism within a single first-stage receptive field. It has been suggested that these contours could be detected by second-stage sign-opponent mechanisms (N. Graham & Sutter, 1998; N. Graham & Wolfson, 2004). Some authors (R. F. Hess & Dakin, 1997, 1999; May & Hess, 2007b, 2008) suggested that ladder detection is based on a second-stage filter fed by first-stage filters orthogonally oriented to the second-stage filter. It is difficult, however, to explain how the response of these mechanisms is affected by the spatial arrangement of the background. Even more difficult is it to account for the phase effect at high-spatial frequencies. There is no obvious reason why phase randomization should improve the response of a sign-opponent mechanism.

A more plausible explanation is provided by models based on spatial interactions between filters with parallel receptive field axes, mediating ladder detection (Yen & Finkel, 1998). These interactions are not only weakly facilitatory (Sarti, Citti, & Petitot, 2009) but also inhibitory (Z. Li, 1998; Polat, 1999), thus accounting for the reduced saliency of ladders with respect to snakes (Yen & Finkel, 1998).

We suggest that inhibition by lateral interactions amongst ladder elements could be reduced by phase randomization. It has been shown that crowding is reduced when target and flankers have different contrast polarity (Chakravarthi & Pelli, 2011). One could assume that there are inhibitory lateral interactions amongst parallel elements that produce a form of crowding and they weaken the link between adjacent contour elements. This interpretation would explain why random phase improves detection, i.e. it may induce a release in reciprocal crowding amongst target elements.

Our results with ladders are hence not trivial, because they seem to refute the most common explanation of ladder contour integration based on the response of a second-stage filter and instead support an interpretation based on lateral interactions between first-order filters.

Combining the findings on snakes and ladders, we can see that a mechanism based on the balance between excitatory and inhibitory lateral interactions may account for the detection of both the contours. Yen and Finkel (1998) proposed an associative field that binds together the adjacent elements with the same orientation, either co-axial (for snakes) or trans-axial (for ladders). The co-axial connections are assumed to be stronger than the trans-axial ones. This model accounts for both the detection of snakes and ladders, but predicts stronger integration for snakes, possibly because of high facilitatory lateral interactions, and less strong integration for ladders, possibly caused by a balance between facilitation and inhibition. A mechanism with these features seems to be the most parsimonious to account for the detection of straight contours.

Chapter 4. Integration of local motion signals: is a cooperative mechanism responsible of dynamic contours?

1. Introduction

Our visual system is optimised to help us navigating in the outside world. To do this, it needs to update information quickly and to extract the most informative cues. For example, it has been showed that the extraction of contours from objects is a fundamental task, since contours contain a great amount of information about the object itself (Attneave, 1954; Feldman & Singh, 2005; Panis & Wagemans, 2009). More importantly, sometimes the system is asked to extract contours of moving objects. How are, then, the static and dynamic information combined together in this task? Several studies have in the past focused on the interactions of static and moving signals (see (Mather et al., 2013) for review). Only few, though, have investigated motion-defined contours (Bex et al., 2001, 2003; Ledgeway & Hess, 2002, 2006; Ledgeway et al., 2005; Nygard, Looy, & Wagemans, 2009). It is intriguing to explore how it this integration happens, i.e. whether form and motion processing operate at the same level and, if they act independently, which one operates first. This issue has been addressed by studies on motion-defined contours. In these studies, observers are usually asked to detect a contour formed by disconnected elements embedded in a background of randomly positioned elements (Field et al., 1993).

For different purposes, both unoriented (Bex et al., 2003; Ledgeway & Hess, 2002; Ledgeway et al., 2005) and oriented (Bex et al., 2001; Ledgeway & Hess, 2006; Ledgeway et al., 2005; Nygard et al., 2009) contour and background patches have been used. Contours made up of unoriented elements are undistinguishable from the random background when static, but they become salient when their elements move in a coherent direction, either orthogonal or collinear to the contour axis (Ledgeway & Hess, 2002; Ledgeway et al., 2005). Pooling of coherently moving signals has been explained according to several models, which take into account velocity (Adelson & Movshon, 1982; Amano et al., 2009; Johnston & Scarfe, 2013), direction (Lorenceau et al., 2005; Vreven & Verghese, 2002) of the signals.

When using oriented drifting Gabors, typical results (Bex et al., 2001; Ledgeway & Hess, 2006; Ledgeway et al., 2005; Nygard et al., 2009) are i) higher saliency of moving with respect to static contours, ii) snakes (i.e. contours made up by collinear elements) are more salient than ladders (contours formed by parallel elements) in both the static ((Dakin & Baruch, 2009; Field et al., 1993; May & Hess, 2007a, 2007b, 2008); see also Chapter 3) and dynamic domain. These results are solid, nevertheless it is not easy to interpret them because both *static* and *dynamic* association fields may use local orientation information. In the case of a static association field, indeed, the prominent role of orientation information is described by the joint constraint of position and orientation (Field et al., 1993). In Chapter 3 it has even been showed that orientation alone is sufficient for detection. If it was the case, motion towards a common direction only represents an additional cue to the presence of a *static* configuration.

However, a dynamic association field is tuned to local orientation of moving signals as well. Whatever the mechanism, indeed, in some cases it has been shown to integrate motion signals in the direction of element configuration (Ledgeway & Hess, 2002; Verghese et al., 2000). Other studies (Casco, Caputo, & Grieco, 2001; Casco, Grieco, Giora, & Martinelli, 2006; Nakayama, Silverman, MacLeod, & Mulligan, 1985; Pavan et al., 2011) found facilitation for contours elements moving orthogonally to their orientation. Another study has even found equal performance in both conditions (Bex et al., 2003).

Studies so far do not give an exhaustive explanation about *what* and *how* gets integrated by the dynamic association field and, in some cases, they even found results hard to conciliate.

In the present Chapter I tried to disentangle one emerging issue: is motion extraction directly involved in contour integration or does it only enhance the salience of static configuration? This question arises from studies showing the “snake superiority effect” (Bex et al., 2001), that would suggest a sort of *supremacy* of the static information. Nonetheless, other studies showed that motion facilitation cannot simply be explained by probability summation (Ledgeway et al., 2005). To do this, I tried to answer a related question, i.e. *what* mechanism underpins dynamic contour integration.

I used a classical path detection paradigm to measure observers' efficiency at segmenting fragmented straight snakes and ladders, either static or moving. I used Gabors micropatterns moving orthogonally to their orientation, by making the carrier drift inside the static Gaussian window. This allows disentangling the role of *static* and *dynamic* association field. Indeed, the static association field operates within the joint constraint of position and orientation, two features of the contour left unchanged when Gabors drift. Comparing static and dynamic condition helps finding specific contributions of a motion system to contour binding. To test for the role of local orientations, both snakes and ladders were used. Finally, I employed straight contours, to limit variance in the range of directions within the contour.

To prove that a dynamic association field exists, I manipulated two parameters expected to affect the response of this mechanism: contour length (experiment 2) and direction of local motion signals (experiment 3). Experiment 1 simply aimed at finding stimulus conditions in which detection of static snakes was not at ceiling and that of ladders was not at floors, so that both could be improved my motion.

2. General method

2.1. Observers

Observers sat in a dark room and were placed at 57 cm from the screen. Viewing was binocular. They were given initial training to familiarize with the stimuli and the task. All subjects had normal or corrected-to-normal visual acuity. Subjects participated voluntarily with no compensation and gave their informed consent prior to their inclusion in the experiment.

2.2. Apparatus and stimuli

Stimuli (Figure 1) were displayed on a 19 inch CTX CRT Trinitron monitor with a refresh rate of 75 Hz and generated with Matlab Psychtoolbox (Brainard, 1997; Pelli, 1997). The screen resolution was 1280 x 1024 pixels. Each pixel subtended ~1.9 arcmin. The mean luminance was 32 cd/m^2. Luminance was measured using a Minolta LS-100 photometer. Stimuli were presented within a square window (11 x 11 deg) placed at the center of the screen. Each stimulus was generated anew immediately prior to its presentation and consisted of a dense spatial array of 121, non-overlapping Gabor patches (see Figure 1). Each Gabor was defined as follows:

$$G(x,y) = m * \sin(2\pi f x x + \varphi) * \exp\left(-\frac{x^2 + y^2}{\sigma^2}\right) \qquad \text{Eq. 1}$$

Where m is the modulation depth of the sinusoidal carrier, f_x is the spatial frequency (SF), φ is the phase of the carrier. The Gaussian is expressed by the exponential part of the equation 1: x and y are, respectively, the horizontal and the vertical distance from the peak of the Gaussian, while σ is the standard deviation (0.21 deg). The contrast was fixed at 0.5 (Michelson contrast).

Gabor patches were placed within a raster of 11 x 11 cells (each cell is 1 x 1 deg) (Field et al., 1993). Each Gabor patch occupied a cell (centre-to-centre distance: 1 deg). The target path was generally made up by five elements (see experiment 1 and 2) with orientation either parallel (snakes) or orthogonal (ladders) to the global orientation of the path (see Procedure section for details). In addition, the path could be either horizontal or vertical. All the Gabors in the path could assume only one orientation value, chosen amongst four: 0 and 180 deg for vertical Gabors, 90 or 270 deg for horizontal Gabors; thus, snakes were defined by elements having the same element and path orientation, whereas ladders were defined by elements with the same orientation but orthogonal to the path's orientation. In addition, 11 deg orientation jitter was applied only to the elements in the target (but see experiment 1). In the dynamic condition (see Procedure section) the elements in the target drifted towards the same direction. The background elements, instead, had a random spatial orientation (ranging from 0 to 360 deg) and position, i.e. a 3 deg spatial jitter was applied to every Gabors in random directions.

Moreover, when dynamic, they were all characterized by random motion direction. All the elements in the array had a phase value set to 45 deg. A set of two relative directions, i.e. 0 deg (motion in a consistent direction always along the contour's axis) and 90 deg (directions orthogonal to the contour) were tested (see Procedure section).

It is important to note that when the local orientations of the drifting 1-d noise were aligned with the depicted contour (0 deg relative orientation), the motion directions were necessarily constrained to be orthogonal to its axis (90 deg relative direction) and vice versa. All the Gabors in the array drifted with the same speed of either 0 or 10 degs^{-1} (see also experiment 4).

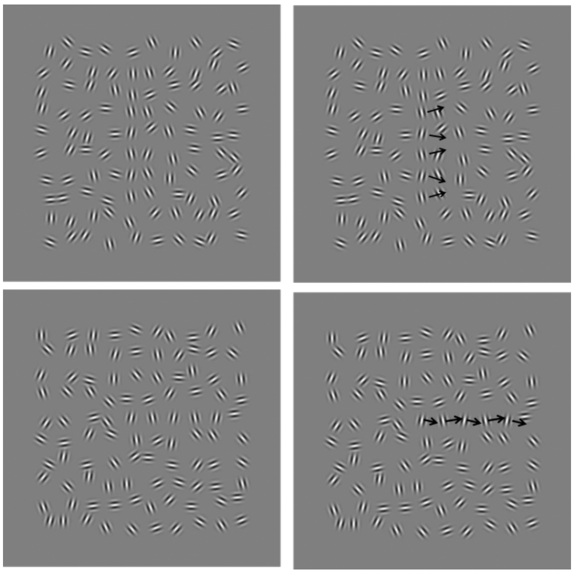


Figure 1. Stimuli and procedure used in the temporal two-interval forced choice task. Top panels show snake displays, whereas bottom panels show ladder displays. Right panels give a representation of the dynamic conditions: black arrows indicate the direction of motion of each element in the contour. Because of drifting, motion was always added in a direction orthogonal to the element local orientation. Contours could be either vertical or horizontal. A vertical snake could move either leftward or rightward, whereas a horizontal one moved always upward or downward. Similarly, a vertical ladder had a upward vs. downward global perceived direction; a horizontal ladder moved either leftward or rightward. Elements in the background were forced to move in random directions, because local orientation could randomly span 360° deg range. See text for further details.

2.3. Procedure

A two interval-forced-choice (2IFC) procedure was used. In each trial, two images were presented (a contour embedded in the background and the background alone) for 200 ms sec each, in random order. Observers were required to choose which of two temporal intervals (Inter-Stimulus-Interval [ISI]: 100 ms) contained the target path. During the ISI, the screen was set to the mean luminance.

Four different target conditions were employed, differing in terms of the presence of motion and of the contour type: (i) static snakes (ii) static ladders; (iii) dynamic snakes, the carriers shifted in a direction orthogonal to the contour axis, i.e., either leftward or rightward for vertical paths, either top or bottom directions for horizontal paths; (iv) dynamic ladders, the carriers shifted in a direction parallel to the contour axis, i.e., either leftward or rightward for horizontal paths, either top or bottom directions for vertical paths.

Immediately after each interval, a masking array (made up by a 11 x 11 Gabor array with elements having random position, orientation, starting phase and, in the dynamic condition, random directions) was displayed.

Path global orientation (i.e., vertical vs. horizontal) and motion direction (i.e., left or right vs. top or bottom) were varied within block. Stimulus type (snakes vs. ladders) and type of motion (static vs. dynamic vs. alternate, see experiment 3 and 4), as well as contour length (experiment 2) were varied across blocks.

The overall percentage correct detection and the associated standard errors (SEs) were calculated separately for each observer and condition.

3. Experiment 1

In the typical contour binding paradigm (Field et al., 1993) straight contours pop out. Since motion is thought to improve contour detection, a facilitation could be negligible on straight contours that are high detectable (ceiling detection performance) even when static ((Field et al., 1993; R. F. Hess & Dakin, 1997, 1999; Ledgeway et al., 2005); see also Chapter 3).

To avoid ceiling effect, following the study by Bex et al. (2001) I introduced orientation jitter (i.e. the relative orientation with respect to the contour axis) along the elements in the path. However, to avoid local confounding effects, the same amount of orientation jitter was applied to all the elements in the path. Finally, a unique value (rather than two distinct thresholds) was used with both contour types. The rationale was that snakes tolerate a bigger perturbation in orientation than ladders. Thus, adding random jitter (as in (Bex et al., 2001)) would generate more noisy local motion signals for snakes than for ladders, making findings not exhaustive.

Experiment 1 was designed to choose the amount of orientation jitter to apply to obtain neither ceiling nor above chance detection performance with both snakes and ladders.

3.1. Method

We compared detection for static snakes and ladders for 5 levels of orientation jitter (1, 11, 21, 31, 41 deg). Orientation jitter was applied clockwise or counter-clockwise, in alternate order and varied between blocks. Ten observers performed at least 10 repetitions per condition (i.e., 80 trials in each block).

3.2. Results

As expected, detection probability decreased as orientation jitter increased for both snakes and ladders. Not to produce ceiling/floor effects, 11 deg orientation jitter was the value chosen to be used in the following experiments for both straight snake and ladder.

4. Experiment 2

In Experiment 2 I ask whether motion extraction intervenes either at local (e.g. micropatterns) level, facilitating local signals integration into a contour, or after contour integration, by enhancing the saliency of the whole structure. Consistently to the first possibility, a plausible candidate for the integration of moving contours might be a cooperative motion system (Chang & Julesz, 1984; Grzywacz et al., 1995; Lorenceau et al., 2005; Snowden & Braddick, 1989; Vreven & Verghese, 2002; D. Williams et al., 1986; Yuille & Grzywacz, 1988) (see also Chapter 2). Cooperative integration, indeed, is a second-order process, in the sense that it integrates local motion signals (Vreven & Verghese, 2002), as the ones used in this paradigm. Differently from other models that focus on direction extraction (Adelson & Movshon, 1982; Amano et al., 2009; Johnston & Scarfe, 2013; Wilson et al., 1992; Wilson & Kim, 1994), it predicts detection of coherent-moving signals embedded in a field of distracters. Thus, it represents a likely candidate for this task.

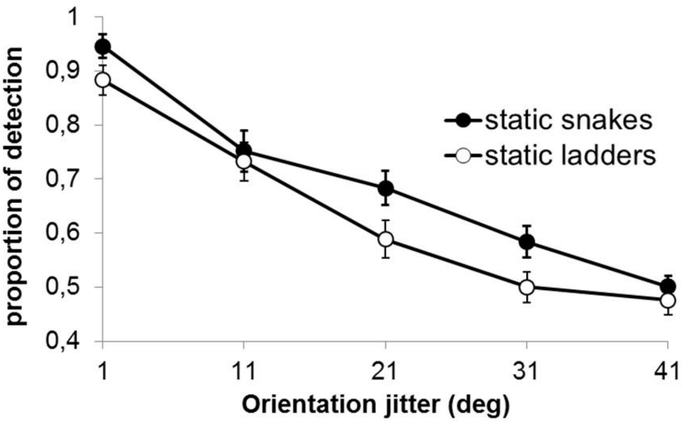


Figure 2. Results from Experiment 1. Observers had to detect a contour (either a snake or a ladder) embedded in a field of randomly oriented distracters. Local orientation of the elements in the path was varied. Proportions of correct answers are plotted as a function of orientation jitter. Solid circles represent snake detection; empty circles are ladder detection. Error bars show the standard error of the mean (SEs).

In experiment 2, we varied the quantity of local signals by changing the number of elements forming the contour. In the static domain, it is well known that detection of static contours improves when increasing the number of elements defining it (W. Li & Gilbert, 2002; Moulden, 1994). This has been interpreted as evidence of cooperative, long-range interactions (Polat, 1999; Polat & Sagi, 1994). Similar results were also found with moving contours (Bex et al., 2003). Indeed, also low-level motion detectors have connections with other detectors in a spatial and temporal neighbourhood. According to the cooperative motion model (Vreven & Verghese, 2002), however, only detectors that are tuned to similar stimulus properties, such as preferred direction, interact with one another. When a single detector is activated by its preferred stimulus, it in turn activates those detectors to which it is connected. Activation from neighbouring detectors with similar tuning, combined with activation from the stimulus, accounts for the enhanced performance for signals with consistent direction (e.g. common fate). Thus, if a cooperative mechanism intervened in the integration of drifting signals, motion facilitation should interact with the number of elements, in a way that the higher is the number of signals provided, the higher is the facilitation by motion obtained. The response of a cooperative motion increases with the increase of the amount of local orthogonal signals (Vreven & Verghese, 2002; Yuille & Grzywacz, 1988).

To test if this is the case, I compared performance at low-signal level (i.e. when the elements in the target are 5) and at high-signal level (i.e. when the contour to spot was 11 elements long). A cooperative motion system would predict higher increase in saliency with longer dynamic contours. I used both snakes and ladders since they are statistically equal, i.e. they contain the same quantity of local motion signals. Moreover, in both cases these signals are orthogonal to local orientation (see General Method section). However, any difference in detection might give further insight about the integration mechanism involved.

4.1. Method

In this experiment, orientation jitter for the contour patches was set to 11 deg for all the conditions tested. We varied the length of the target, i.e. the number of elements forming it, of 5 levels ranging from 3 to 11. The number of elements in the contour, as well as contour orientation (vertical vs. horizontal) and element direction (left vs.

right), were varied within blocks. Ten naïve subjects performed at least 5 repetitions per condition (i.e., 200 trials in each block).

4.2. Results and discussion

Figure 3 shows results from this experiment, for snakes (solid circles, figure 3) and ladders (empty circles, figure 3) separately for static (figure 3a) and dynamic condition (figure 3b). Again, we plotted proportion of correct answers, this time, as a function of contour length. Focus now is low- and high-level amount of signals, therefore statistics were conducted on short (3 elements) and long (11 elements) contour conditions only. From a quick look at figure 3, one can easily see that snakes are generally more detected than ladders. Moreover, a different effect of contour length appears in the static and dynamic conditions.

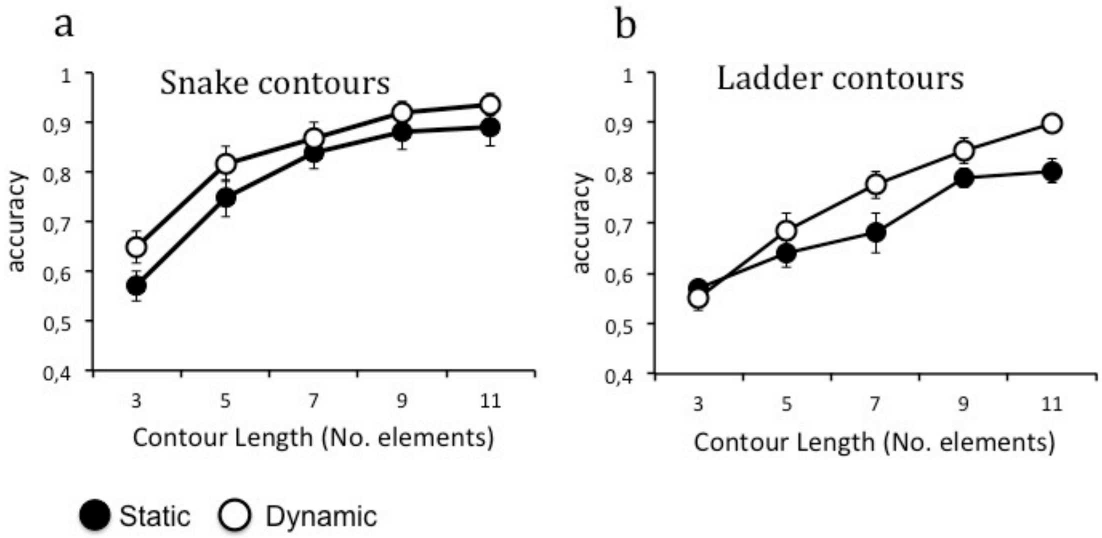


Figure 3. Proportion of correct answers as a function of contour length for snake contours (a), and ladder contours (b) separately for snakes (solid circles) and ladders (empty circles). Error bars display SEs.

A two-way repeated measures ANOVA with Stimulus Type (snake vs. ladder), Motion (static vs. dynamic) and Contour Length (3 vs. 11) as main factors revealed a significant effect of Motion ($F_{(1,8)}$=10.467, p=0.012) but not of Stimulus Type (p=0.087), indicating that moving contours led to easier detection static contours in general. We also found a significant effect of Contour Length ($F_{(1,8)}$= 202.124, p<0.0001). Therefore, motion and contour length are two factors that facilitate contour integration. More interestingly, the interactions between Motion and Contour Length ($F_{(1,8)}$=9.949, p=0.014) and Motion x

Contour Length x Stimulus Type ($F_{(1,8)}$=7.868, p=0.023) resulted significant. Let's analyse the three-way interaction first. Bonferroni-corrected pairwise comparisons revealed that snake detection was generally facilitated by motion at both contour lengths (p=0.009 and p=0.038 with 3 and 11 elements, respectively). Ladders, instead, were facilitated at 11-elements length only (p=0.013 vs. p=0.440 with 3 elements), probably due to the generally low salience of this stimulus. The two-way interactions showed, more generally, that motion helps detection, but facilitation is only significantly visible at longer lengths (p=0.002 vs. p=0.136 with 3 elements). No other interaction was significant.

These results are in agreement with our predictions: adding motion produces a significant increase in detection. The improvement, moreover, is not equally present for all contour length, but increases (and becomes significant) as the number of *motion* signals increases. This shows that motion does not simply have an additive effect on detection. Results are thus consistent with the idea of a cooperative motion system.

Finally, it is worth pointing out that, due to the presence of orientation jitter, these results are likely to reflect contour integration and do not simply show averaging of same common direction in a field or randomly directed elements. Local motion signals are indeed integrated into a global percept in spite of small variations in orientation and direction.

To summarise, experiment 2 tested the mechanism by facilitating integration. Experiment 3 will further assess the topic by trying to disrupt integration.

5. Experiment 3

In Experiment 3 we further sought evidence for a dynamic association field that allows segmentation of the contour path by combining local motion signals with similar directions. To this end we perturbed the relative direction of motion of the elements composing the path, by changing it of 180 deg, in alternation. In our prediction, it is a cooperative mechanism that enhances path saliency by detecting a pool of local motion signals drifting in the *same direction* (Chang & Julesz, 1984; Grzywacz et al., 1995; Lorenceau et al., 2005; Snowden & Braddick, 1989; Vreven & Verghese, 2002; D. Williams et al., 1986; Yuille & Grzywacz, 1988). Perturbing (i.e., alternating) direction should disrupt common fate and prevent grouping by a cooperative motion mechanism, thus worsening contour detection, when comparing it to coherent-direction condition.

5.1. Method

In this experiment we replicated across blocks the static and dynamic conditions already tested in previous experiments plus an alternate motion condition. In this condition, the direction of the first element in the path was chosen randomly, the subsequent was rotated by 180 deg and so on. For example, if the first element was moving leftward, the second was drifting rightward, the third leftward etc. Orientation jitter for the contour patches was set to 11 deg for all the conditions tested. The target was five elements long. One author and seven naïve subjects participated to Experiment 3. Observers performed at least 5 repetitions per condition (i.e., 40 trials in each block).

5.2. Results and discussion

Overall, results (Figure 4) show that static stimuli were less salient than coherent direction contours and more salient than alternate direction contours. This happened for both snakes and ladders.

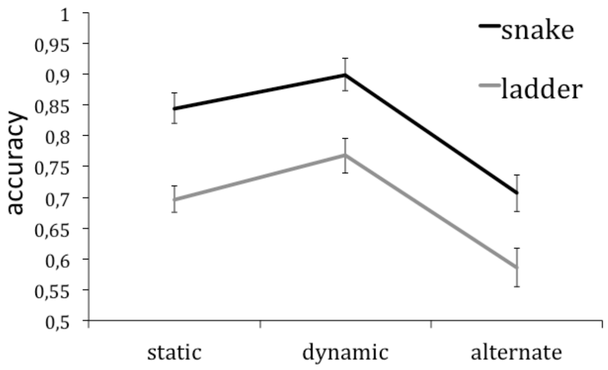


Figure 4. Results are proportion of correct answers as a function of motion condition: static contours vs. same direction vs. alternate direction. The black line shows performance with snakes. The grey line shows performance with ladders. Error bars are the standard error of the mean (SEs).

This was confirmed by a two-way repeated measure ANOVA with Contour Type (snakes vs. ladders) and Motion (static vs. dynamic vs. alternate) as main factors. It pointed out a significant effect of Contour Type ($F_{(1,7)}$=114.231, p=0.0001) and of Motion

($F_{(1,7)}$=40.736, p=0.0001), no significant interactions were found. Pairwise comparisons (Bonferroni corrected) revealed significant differences between static and dynamic (p=0.012), static and alternate (p=0.001) and dynamic and alternate (p=0.001) conditions, demonstrating that contours with alternated direction were less detected than in the other moving conditions.

These findings were consistent with the idea that signals moving in a coherent direction are pooled together and give further support to the idea that the extraction of motion signals interact with that of static cues. Moreover, the fact that alternating direction determines such a detrimental effect implicates that dynamic signals are extracted prior to static signals. I will come back to this in the Discussion.

6. Experiment 4

Experiment 3 pointed out a large effect by motion direction: contour made up by elements drifting towards opposite direction are even less detectable than static ones. Results are consistent with those by Hayes (2000). Differently from this experiment, the author perturbed the relative static position of the elements in the contour. He made observers judge the alignment of a moving contour; one of the elements was positionally shifted. He found that observers perceived misaligned Gabor to be aligned when drifting in a direction opposed to the drift. According to the author, this was due to the fact that our visual system integrates contour on the basis of their *perceived* - rather than their *physical*- position, according to the well-known finding that drifting objects are seen shifted towards motion direction (motion induced positional shift (Chung, Patel, Bedell, & Yilmaz, 2007; De Valois & De Valois, 1991; Whitney, 2002)).

Thus, also positional shift and misaligned perception represent a satisfactory interpretation of the results of experiment 3.

In the static domain, it is well established that contour binding relies on a joined constraint of position and orientation (Field et al., 1993). Chapter 3 have proved the importance of local orientation, as well as previous findings by Casco et al. (2009). Some authors have instead stressed the prominent role of position (R. Watt et al., 2008). With dynamic contours, less is known about the relative importance of position and orientation. Ledgeway and Hess (2006) demonstrated that coherent orientation provides better detection performance. One may argue that motion integration might be influence by similar principles to those underlining the static association field.

To test this hypothesis, in experiment 4 we increasingly perturbed the relative perceived position of the elements forming the target by varying speed. Indeed, the perceived shift in position increases as speed increases (Bex et al., 2001; Chung et al., 2007; Whitney, 2002). If position information was taken into account, then I should expect worse detection performance when elements drift at high speed and this should occur in the alternate direction condition *only*.

6.1. Method

I employed the same method as in experiment 3. I tested cotour detection in two motion conditions (coherent vs. alternate direction). Only snakes were used in this condition. I also used three speed levels (5, 10 and 20 degs^{-1}) that were varied between blocks. Orientation jitter for the contour patches was set to 11 deg for all the conditions tested. The target was five elements long. One author and six naïve subjects performed at least 5 repetitions per condition (i.e., 40 trials in each block).

6.2. Results and Discussion

Our findings (Figure 5) showed that performance in the alternated direction condition was worsened by speed, whereas it remained constant in the coherent direction condition.

A two-way repeated measure ANOVA with Motion (coherent vs. alternated direction) and Speed (5 vs. 10 vs. 17 degs^{-1}) as main factors pointed out a significant interaction (Motion x Speed, $F_{(2,12)}=13.875$, $p=0.001$), together with Motion main effect ($F_{(1,6)}=12.667$, $p=0.012$). The difference between the two motion conditions was significant only at the higher speed levels, 10 degs^{-1} ($p=0.003$) and 17 degs^{-1} ($p=0.005$; pairwise comparisons Bonferroni corrected).

This might indicate that, at low speed, perceived positional shift is weaker and thus it does not differentiate the performance in the two conditions. As speed increases, elements in the contour are perceived as more misaligned and performance decreases. Perceived position therefore is a constraint that influences moving signal integration, at least with snakes.

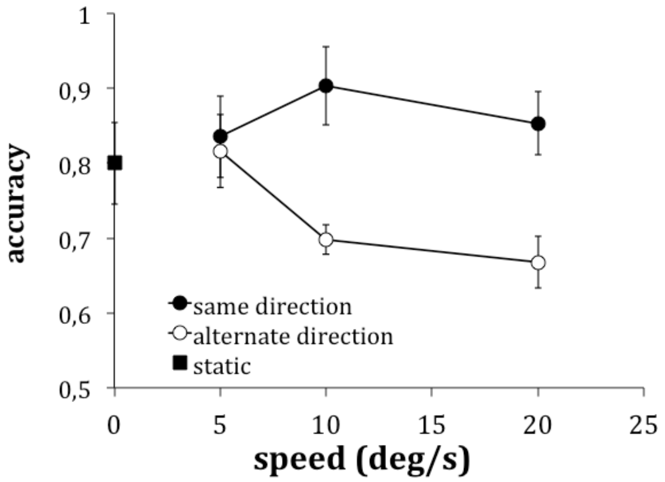


Figure 5. Results are proportion of correct answers as a function of elements speed, for two motion conditions: same direction (solid circles) vs. alternate direction (empty circles). The solid square shows performance with static stimuli. Error bars are the standard error of the mean (SEs).

7. General Discussion

In the present Chapter, I investigated the integration of dynamic contours made up of Gabor micropatches. I varied the number and the direction of the elements in the contour and I made micropatches drift at different speeds. To summarise, I generally found a higher detectability of snakes and of moving contours; motion facilitation specifically occurs when the number of *motion* signals increases. More interestingly, I proved that alternating the direction of the elements in the target significantly worsened the performance. Importantly, this is true not only with respect to the dynamic condition, as one would expect, but most of all with respect to the static condition. This effect was even more evident when the drifting speed was higher.

These results confirm previous findings (Bex et al., 2001; Ledgeway & Hess, 2006; Ledgeway et al., 2005; Nygard et al., 2009) that contour detection is facilitated by motion. What is new, however, is the evidence of an active influence of motion on static signals. Indeed, alternating the direction of the Gabor patches reduce performance with respect to the baseline performance (measured in the static condition). This is intriguing because the alternating direction should affect the motion domain and leave the static domain unchanged. How to explain then the direction effect? The answer is suggested by Hayes (2000) when he considers the positional shift illusion induced by motion on a drifting Gabor carrier (Chung et al., 2007; De Valois & De Valois, 1991;

Whitney, 2002). He made observers judge the alignment of a moving contour; one of the elements was positionally shifted. He found that observers perceived misaligned Gabor to be aligned when drifting in a direction opposed to the drift. This leads to the suggestion that the positional information used in contour integration is the perceived (non-veridical) one, rather than the physical one. To this extent motion perturbs a *static* attribute of the stimulus and snakes are no longer visible because shift gets as big as speed increases. The consequence is a non-optimal activation of the *static* association field, given that the orientation-position constrained is no longer respected. Thus, according to our data, motion seems to intervene and modulate static association field operations.

If this were the case, the obvious prediction is that the effect of alternating motion shouldn't apply when elements have opposite direction but are not perceived displaced from the contour backbone. To test for this prediction, we replicated experiment 4 using ladder contours. Results in figure 6 show that increasing speed does not affect ladder detection, an effect compatible with a lack of perceived displacement from the contour backbone (however, note that results in this condition are generally lower than in the coherent-direction condition).

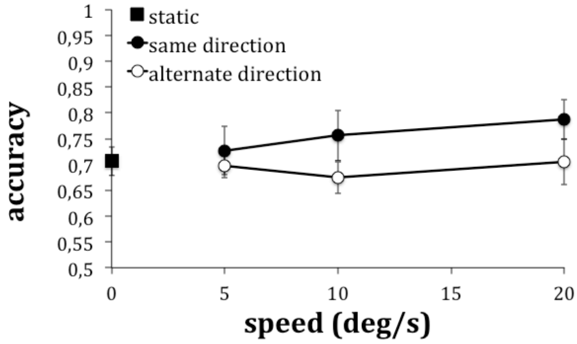


Figure 6. A ladder detection task with two motion conditions (coherent vs. alternate direction). Three speed levels (5, 10 and 20 $degs^{-1}$) were varied between blocks. Orientation jitter for the contour patches was set to 11 deg for all the conditions tested. The target was five elements long. One author and seven naïve subjects performed at least 5 repetitions per condition (i.e., 40 trials in each block). A two-way repeated measure ANOVA reported a main effect of Direction ($F_{(1,7)}$=6.287, p=0.041), with alternate direction worsening the performance. No main effect of Speed ($F_{(2,14)}$=1.303, p=0.303) and no interaction ($F_{(2,14)}$=0.772, p=0.481) were found. Results are proportion of correct answers as a function of elements speed, for two motion conditions (solid and empty symbols). Error bars show the standard error of the mean (SEs).

In apparent contradiction with our results, Bex et al. (2001) found that performance in contour detection was not influenced when elements in the path drifted at different speeds (spanning as much as three octaves). However, although motion caused perceived displacement, common fate was preserved in this condition. The authors also concluded that the facilitatory effect of motion is based on the common direction of motion of the elements, rather than their common speed, temporal frequency or phase. Common direction allows integration over a wide range of speed. Taken together, these results, together with those from experiment 3 and 4, further suggest that detection of moving contours is mainly based on direction. A similar consideration arises from the finding that contours made up of flickering elements, i.e. having ambiguous direction, do not show motion facilitation (Bex et al., 2001; Ledgeway & Hess, 2006). In addition, also Ledgeway and Hess (2002) found that direction is a powerful cue, and it is even more strong if compared to speed (R. F. Hess & Ledgeway, 2003), for detecting motion defined contours. Finally, the importance of common direction was also found with more complex-contour stimuli (Rainville & Wilson, 2005).

The importance of local direction suggests that a cooperative motion mechanism integrates coherently moving local motion signals (Vreven & Verghese, 2002). This applies to both snakes and ladders, because they are both tuned to common direction of the elements forming them. This is more evident for ladders, because with these stimuli we showed an effect of motion for suprathreshold stimuli (i.e. motion incremented performance with long contours; see experiment 2). Results are however consistent for snakes, even if results from experiment 2 are not significant. This may be due to the high saliency these contours classically have (Bex et al., 2001; Dakin & Baruch, 2009; Field et al., 1993; Ledgeway & Hess, 2006; Ledgeway et al., 2005; May & Hess, 2007a, 2007b, 2008).

It may be objected that the two contour types are not symmetric as regard as motion information they convey: indeed whereas in snakes both local and global motion are orthogonal to the contour, this is not the case for ladders that have only local motion orthogonal to the contour. However, a similar effect of alternated motion for the two stimuli suggests that this difference does not matter. This might suggest that the cooperative mechanism involved is a “local” process. Other evidence in support for these come to the fact that integration of moving contours was found to be tuned to spatial frequency and local orientation (Bex & Dakin, 2002; Ledgeway & Hess, 2006).

Not in contradiction is the fact that the mechanism described by Vreven and Verghese (2002) is a second-stage mechanism, because “second-stage” refers to the fact that it describes interactions between the responses of low-level motion detectors. Finally, moving contours have also found to be indifferent to spatial frequency and contrast polarity (Bex et al., 2003), as usually observed with global motion detection (Yang & Blake, 1994). However, although the authors extend these results to the contour integration domain, it should be noted that in this case elements composing the contour actually moved across the display, thus introducing a confounding effect.

Last but not least, I provided evidence for interplay between static and dynamic cues (see results from experiment 3). What are the properties of this interplay? My suggestion is that motion signals are extracted via a cooperative mechanism *prior* to static signals integration. When signals drift coherently, even if an illusory position shift occurs, the whole configuration is moved, thus leaving relative positions and orientations unaffected. At the same time, the cooperative motion enhances links within those elements coherently moving, helping the static association field and thus explaining the facilitation in detection that occurs with motion. On the other hand, alternate direction disrupts the relative position of the elements on the contour backbone. The input provided to the static association field is now non optimal, since the joined constraint of position and orientation is perturbed. Performance is then worse in this than in the static condition. In line with these findings is the fact a contour configuration (e.g. a series of patches disposed to form a S-shaped contour) undetectable when static becomes visible when moving (Ledgeway & Hess, 2002; Ledgeway et al., 2005).

To sum up, I have showed the presence of a cooperative mechanism that integrates local moving signals sharing a common direction. Motion extraction occurs prior to static integration and determines whether this will be successfully or unsuccessfully performed. This happens because motion extraction enhances the presence of a contour configuration thus facilitating detection by the association field.
The study in Chapter 5 will further try to shed light on this issue.

Chapter 5. Electrophysiological correlates of moving contours.

1. Introduction and aim of the study

Despite what early works in vision science suggest (Mishkin & Ungerleider, 1982; Mishkin et al., 1983; Ungerleider & Pasternak, 2004), motion and form analysis constantly interact in the brain. A growing number of data have challenged the idea that the visual system is anatomically and functionally segregated into "what" and "where" streams (Ellison & Cowey, 2006, 2007, 2009; Mather et al., 2012; Pavan et al., 2013; Wokke, Scholte, & Lamme, 2014). The so-called "what" information modulates activity not only in ventral, but also in dorsal areas (and vice-versa) (Braddick, O'Brien, Wattam-Bell, Atkinson, & Turner, 2000; Hesselmann & Malach, 2011; Konen & Kastner, 2008). Besides, studies on animals have revealed reciprocal inter-stream connections at all levels of the visual hierarchy (Felleman & Van Essen, 1991; Van Essen & Maunsell, 1983). This new perspective has recently known an increasing success ((Mather et al., 2013) for a review). However how and where this interaction takes place is still debated.

Wokke et al. (2014) directly attempted to assess contributions from ventral and dorsal stream in a motion-defined figure segmentation task by combining EEG recording with TMS. They found that disrupting V5 activity impaired the task, whereas disrupting LO enhanced performance. Both effects were visible when TMS was applied in a short time window (i.e. 100 ms) after stimulus onset. Authors concluded that the two systems compete in this task. However, it has also been shown that areas in dorsal and ventral system cooperate to optimally perform a task (Ellison & Cowey, 2009).

A powerful way to test motion-form interaction might be to investigate the integration of moving contours. Much is known about the integration of static signals into contours. One influential hypothesis asserts that an association field groups together signals following the joint constraint of position and orientation (Field et al., 1993), although also position (R. Watt et al., 2008) and local orientation (see chapter 3) alone can explain detection. To date, the topic has been intensively studied by means of a contour integration paradigm, in which observers are asked to detect a Gaborised contour embedded in a field of random distracters (see chapters 2, 3 and 4).

Other studies have instead focused on the integration of moving contours, usually showing higher detection when moving signals move non rigidly, i.e. one towards the following one (Ledgeway & Hess, 2002; Vreven & Verghese, 2002) (but see (Bex et al., 2003)).

Finally, only few studies combined both static and dynamic signals and studied the integration of contours defined by orientation and motion. These studies, as the one reported in chapter 4, unveil interplay between the two factors. Indeed, snake contours, which are more easily detected than ladders when static, are also more salient when dynamic (Bex et al., 2001; Ledgeway & Hess, 2006; Ledgeway et al., 2005). More importantly, motion enhances detection and common direction determines which elements are going to be perceived as part of the target and what are going to be considered as distracters (Bex et al., 2001; Hayes, 2000; Ledgeway & Hess, 2006; Ledgeway et al., 2005; Nygard et al., 2009). This can be explained in a simple way, asserting that motion simply increases saliency. On the other hand, results in Chapter 4 question this option and support the idea that motion signals are integrated *before* orientation signals. Perturbing the relative direction of the elements in a contour was found to worsen performance with respect to both the dynamic (same direction) and static baseline. I therefore concluded that motion perturbation (at least) interacts and deteriorates the extraction of static signals prior to their integration, i.e. the mechanism responsible for moving contour integration should evaluate dynamic prior to static signals.

There is a general agreement assuming that the integration of moving and orientation signals might be carried on by two distinct mechanisms (Bex et al., 2001; R. F. Hess et al., 2003; Ledgeway & Hess, 2006; Ledgeway et al., 2005). In Chapter 4 I have showed that direction, together with position, are the main features *feeding* the dynamic association field. Moreover, such dynamic association field differs from the static one also because moving contours are narrowly tuned to both spatial frequency and local orientation (Ledgeway & Hess, 2006), whereas static contours can tolerate larger differences in spatial frequency (Dakin & Hess, 1998). Therefore, putting together results from Chapter 4 and (Bex et al., 2001; Ledgeway & Hess, 2006; Ledgeway et al., 2005), one might suggest that there are two distinct mechanisms for static and dynamic integration, possibly with the dynamic one acting earlier.

Two questions arise. The first one is "how" the difference between the two fields unfolds. The second is clearly about the dynamics (and consequently neural substrates) involved. One might try to answer this question helping psychophysics with other techniques, such as visual evoked potentials recording.

Previous studies on contour integration and evoked potentials showed that the presence of a contour elicits an early negativity (N1, (Casco et al., 2009; Machilsen et al., 2011; Mathes et al., 2006)). Machilsen et al. (2011) found that N1 (a negative peak within 145 and 250 ms) only emerged when a contour was displayed, whereas background alone did not enhanced negativity in this temporal window. Differences in positivity were instead found when the local orientations of the background elements were manipulated. Other authors (Shpaner et al., 2013) confirmed these results and also proved that the presence of a contour is not related to variations in amplitude of earlier components, such as C1.

This early negativity is so traditionally related to the presence of texture borders (Bach & Meigen, 1992, 1997; Caputo & Casco, 1999; Casco et al., 2009), that authors have proposed that texture segmentation and contour integration might lie on the same mechanism (Machilsen et al., 2011). Bach and Meigen (1997) compared parameters of the segmentation component in four conditions, i.e. luminance- vs. orientation- vs. motion- vs. disparity-defined textures. They found that the components associated to different texture dimensions were surprisingly similar, peaking at the same latency and showing more or less the same amplitude. They concluded that the mechanism responsible of segmentation generally retrieves gradients in a visual scene, thus ignoring low-level visual dimensions. It seems, then, that an non-specific segmentation process could characterise both association fields, at least at early levels. A difference between the two might arise later. In agreement, Schoenfeld et al. (2003) found that both motion- and luminance- defined textures produced an early negativity, whereas motion-defined stimuli produced slower activity (60 ms delay) compared to luminance-defined textures over the lateral occipital and posterior temporal scalp areas. The authors concluded that motion stimuli need elaboration from an additional area in figure perception (i.e. MT+), whereas luminance-defined figures are directly elaborated by higher areas in visual hierarchy. Also Wokke, Sligte, Steven Scholte, and Lamme (2012) observed negativity (137-211 ms) associated to the simple presence of a motion-defined border. They further found that early TMS on V1/V2 reduced border detection. Other findings even identify (late) additional areas, specifically recruited in processing moving contours. In an fMRI study, Orban et al. (1995) found a selective activation of a late stage in motion processing, i.e. kinetic occipital region (KO), in response to motion-defined texture borders, whereas area MT was found to generally respond to motion, either uniform or creating contrast borders.

Other researchers believe instead that V1 underpins both orientation- and motion-defined contours (Lamme, van Dijk, & Spekreijse, 1993, 1994). Consistently, some cells in V1 have been found to respond to feature contrast per se, either when contrast was generated by elements having different local orientation or moving in the opposite direction (Kastner, Nothdurft, & Pigarev, 1999).

Nonetheless, other authors have pointed out the importance of recurrent feedback processes to early cortical areas (Fahrenfort, Scholte, & Lamme, 2007; Foxe & Simpson, 2002; Lamme, Zipser, & Spekreijse, 2002; Robol, Grassi, & Casco, 2013) and some of them have identified them as the main cause of activation found in V1 during contour integration tasks (Shpaner et al., 2013). Contour integration, in this theoretical framework, is a high-level task mainly performed by visual areas, such as V4 and LOC, that send feedbacks to lower level areas, such as V1 (Shpaner et al., 2013). In agreement with contour integration being a high-level process are also data by (Tanskanen et al., 2008).

To summarise, studies so far have showed a great number of areas involved in contour integration and moving contour integration. Some studies report a prominent role of V1. Others believe that V1 is modulated by feedback connections and point out motion-contour tuned processing in late areas. Still, they do not help disentangling my question, i.e. how it is that motion facilitates contour integration. In the present work, I directly compared the temporal dynamics associated to orientation-only and orientation+motion defined contours with visual event-related potentials (VEPs).

Casco et al. (2009) suggested that two processes, with two separable electrophysiological correlates, could underpin contour integration. They indeed isolated an early component, (40-179 ms) that increased in positivity with *grouping*, and a late component (275 ms) that increased in negativity with *segmentation*. To this extent, correlates of motion facilitation can modulate either a grouping or a segmentation component, when compared to static integration. To specifically asses the modulation of early components in this task, I used a design explicitly apt to elicit C1 (Shpaner et al., 2013). The design was originally proposed by (Shpaner et al., 2013) and consisted in presenting stimulus displays in different quadrants of the visual field. There is evidence indeed that C1 is strongly retinotopic, thus different positions elicit different polarities of this component (Clark et al., 1994; Jeffreys & Axford, 1972). Following Casco et al. (2009), I predict that an early influence by motion should affect the grouping component, whereas late influence should be observable from segmentation.

2. Method

2.1 Observers

The author and 14 observers (mean age: 24, std: 2.83) participated to the experiment. The observers were naïve to the purpose, provided written informed consent prior to their inclusion in the experiment and participated voluntarily with no compensation. Observers sat at 70 cm from the monitor and were asked to limit movements and eye-movements to avoid artefacts in VEPs recording. Viewing was binocular. All subjects had normal or corrected-to-normal visual acuity. All procedures were approved by the ethical board of the School of Psychology (University of Padua) and data were collected conformed to the Declaration of Helsinki.

2.2 Stimuli

Stimuli were displayed on a CTX CRT with a refresh rate of 60 Hz and a mean luminance of 35 cd/m^2 (measured with a Minolta LS-100 photometer). Stimuli were generated via Matlab and elements of the Psychtoolbox (Brainard, 1997; Pelli, 1997). The screen resolution was 1024 × 768 pixels. Each pixel subtended ~1.8 arcmin.

Stimuli are showed in Figure 1a. They were presented within a square window (10 × 10°) placed either to the upper left or the lower right quadrant of the screen. Off-center presentation was necessary to evoke the retinotopically generated C1 component (Clark et al., 1994; Jeffreys & Axford, 1972).

Each stimulus was generated anew immediately prior to its presentation and consisted on average of a dense spatial array of ~96 (std: 2) Gabor patches. Each Gabor was sinewave, had a spatial frequency of 3 cycles/° and standard deviation 7.2 arcmin. The size of each patch was 45.5 arcmin. Contrast was set to maximum.

Stimuli were created as follows: in contour-displays, contours were generated as a vertical straight contour of 7 elements passing through the centre of the stimulus array. Centre-to-centre distance was set to 44.9 arcmin. Contours were embedded within a background of randomly oriented distracter elements. To do this I placed distracter elements using an iterative procedure that selected random (x,y) coordinates within the array, ensuring a minimum mutual centre-to-centre distance of 72 arcmin between all elements in the display. This procedure produces more uniform element density near the contour (Dakin & Baruch, 2009) than the classic one introduced by Field et al. (1993), thus minimising perceptual and density differences between contour- and no contour- displays.

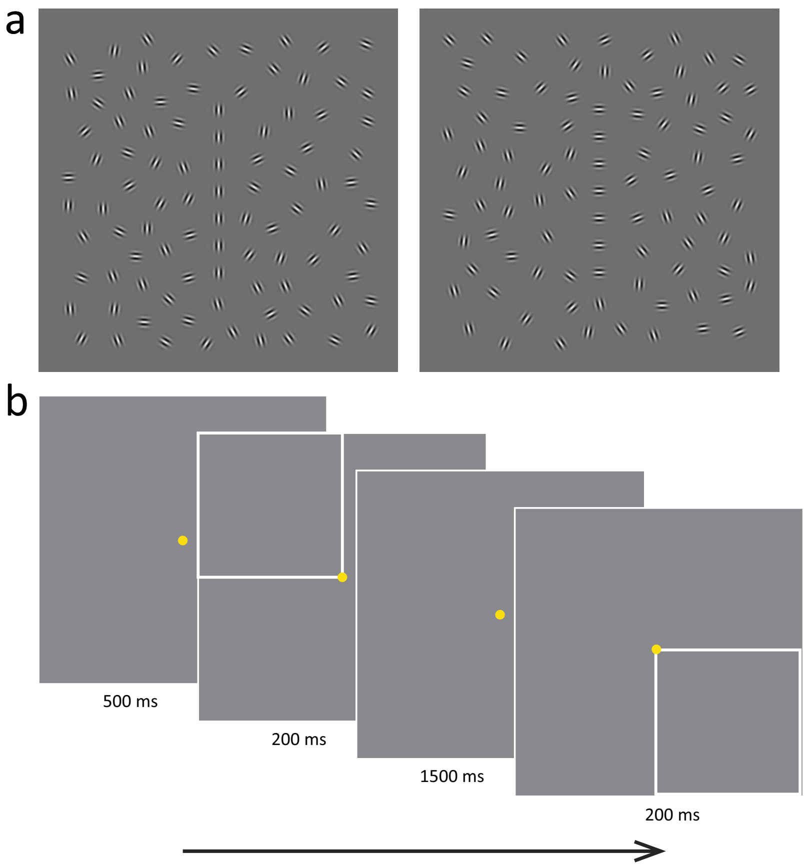


Figure 1. Stimuli and procedure used in the temporal two-interval forced choice task. (a) The panels show respectively a snake –left- and a ladder –right- contour embedded in a field of randomly positioned and oriented Gabors. Contours were straight and formed by seven vertically aligned iso-oriented elements. (b) Observers fixated at a central point for 500 ms. The first interval was then displayed for 200 ms. During the inter-stimulus interval (1500 ms) the fixation point was showed on a grey display. The second interval was showed for 200 ms. Observers had to indicate whether the target was presented either in the first or in the second interval. For explicative purposes white frames are displayed instead of Gabor arrays.

No-contour displays were generated according to the same iterative procedure, with the exception that no contour locations were excluded. Elements in the target had a common local orientation of either 90 (*snakes*) or 0 (*ladders*) deg, whereas each distracter had a local orientation randomly selected in a range spanning 360 deg.

In the dynamic condition (see Procedure) the carriers of the Gabors in the array drifted behind the envelope apertures with a speed of 4 $degs^{-1}$. Since direction of drifting is constrained by the carrier local orientation, all the elements in the target moved in the same direction (rightwards for snakes, downwards for ladders), whereas distracters had random directions.

2.3 Procedure

Observers performed a temporal two-alternative forced choice task (Figure 1b), in which they were asked to indicate which interval contained the target, either the first or the second. They were instructed to fixate at a central point (a yellow circle with diameter 14.4 arcmin) presented for 200 ms. The first interval was displayed immediately after for 200 ms and followed by an inter-stimulus interval of 1500 ms, during which the monitor was set to the minimum luminance. The second interval was then display for 200 ms. After an inter-trial interval (3000 ms) a new stimulus was displayed. Note that stimuli could occupy either the upper-left or the lower-right quadrant of the monitor. At every trial, the display position was randomly chosen in a way that if the first interval was displayed upper-left, the second was constraint to occupy the lower-right. The presence of the target in the first or second interval and the position of the first and the second interval were balanced within a block.

We collected data in two separate conditions: i) static integration; the stimuli were static; ii) dynamic integration; the stimuli were composed by drifting Gabors. The execution order was balanced between observers, 7 starting with the static condition and 7 performing the dynamic condition first. To allow participants to familiarise with the task, recording was preceded by a short training session (10 trials per condition). Participants starting with the static condition were trained with static stimuli and vice-versa.

Stimulus type (snakes vs. ladders) and target position (upper-left vs. lower-right) within block. Each condition was performed 80 times, thus each block was composed by 320 trials. Each block had a mean duration of 40 min.

2.4 VEPs recording

Continuous electroencephalographic activity (EEG) was recorded from 30 scalp electrodes (FP1, FP2, Fz, F3, F4, FC1, FC2, FC5, FC6, Cz, CPz, CP1, CP2, CP5, CP6, Pz, P7, P5, P3, P4, P6, P8, POz, PO7, PO3, PO4, PO8, Oz, O1, O2) using sintered Ag/AgCl ring scalp electrodes and Brain-Cap, labeled according to the 10–20 international system. All scalp channels were referenced to the average reference during recording, because with the QuickAmp72 the uni-polar electrophysiological inputs are configured as a reference amplifier. The ground electrode was positioned in front of Fz (Af). The EEG was amplified, band-passed (0.1–100 Hz), and digitized at a sampling rate of 500 Hz (Recorder software, QuickAmp amplifier). A Parks-McClellan notch filter centered on 50 Hz was also applied. Scalp electrode impedance was maintained below 10 kΩ. Scalp electrooculogram (EOG) was also recorded bipolarly through four additional electrodes placed left and right of external canthi for horizontal eye movements, and above and below the right eye for blinks and vertical eye movements. All trials in which the observer made an eye movement larger than 1° were rejected.

2.5 Data analysis

EEG and ERPs derivations and analysis were performed using the Eeglab toolbox (Delorme & Makeig, 2004) and the Erplab plugin for Eeglab (Lopez-Calderon & Luck, 2010) for Matlab. EEG recording from each channel were offline filtered with a 45 Hz low-pass filter. For each participant, at each trial, the continuous EEG was segmented into two 700 ms-long epochs that were timelocked to the stimulus events (from −200 ms before stimulus onset to 500 ms after stimulus onset). This way, one epoch was associated to one stimulus interval and this allowed comparisons between epochs sharing one feature but differing from the others (in other words, we could compare components associated to snakes vs. ladders when the stimulus was upper-left displayed, and so on). Trials associated to response errors and/or contaminated by artefacts (eye blinks, eye movements, or muscle potentials) at any electrode were excluded from the average. On average 10 % of the trials were excluded due to presence of artefacts. I used an artefact rejection criterion of ±100 μV, to reject trials with eye movements, blinks and other noise transients. Prior to group-averaging, ERP data were baseline corrected from −50 to +20 ms. Individual subject averages were then generated for each of the four target condition (snake vs. ladders and up-left vs. down-right) and for the distracter conditions (up-left vs. down-right). Difference waves (D waves) were computed: we subtracted ERPs associated to distracter displays from ERPs obtained with target display. D-waves thus reflect the activity associated to contour detection, without the activity linked to local orientations and single elements processing.

Based on previous studies on figure-ground segregation and contour interactions (Casco et al., 2009; Machilsen et al., 2011; Pitts, Martinez, Brewer, & Hillyard, 2011; Scholte, Jolij, Fahrenfort, & Lamme, 2008; Wokke et al., 2012), we restricted our interest to a subset of channels (POz, PO7, PO3, PO4, PO8, Oz, O1, O2). However, in this preliminary report, only D-waves from Oz, O2 and O1 were analysed. The D-waves were characterized by a series of components with latency ranging from 40 ms to 220 ms (see Table 2). P1 was chosen as the most positive component with a latency spanning 80-140 ms. Given its inter-observers variability, the earliest component (C1) was then identified on individual basis as the first highest deflection peaking before P1 in a time window from 40 to 100 ms (Kelly et al., 2008). As in (Casco et al., 2009; Nikolaev & van Leeuwen, 2004; Rossi & Pourtois, 2012), I chose the earliest negative component (i.e. the one that reflected more negativity in the contour+background condition). N1 was the most negative component emerging after P1 at 100-200 ms. I found that varying the position of the stimulus on the visual field modulated C1 (as expected (Di Russo et al., 2002; Kelly et al., 2008; Shpaner et al., 2013)), but also caused polarity-inversions of P1 and N1. The polarity-inversion effect is clearly visible from figure 2. This was also found in a co-registration fMRI-ERPs study by Di Russo et al. (2003) and it was thought to mirror delayed, re-entrant input into V1 from higher visual areas. Re-entrant inputs might enhance figure/ ground segregation and improve the selection of relevant from irrelevant stimuli at attended locations in the visual field (Lamme & Spekreijse, 2000; Roelfsema, Lamme, & Spekreijse, 1998).

Comparisons in this preliminary study were conducted on C1 and N1 components, because there is evidence that P1 is not modulated by the contour (Casco et al., 2009; Machilsen et al., 2011).

Since I was interested in assessing amplitudes and latency modulation of N1 (the so-called segmentation component), from now on figures showing components related to the upper-left conditions will show negative values on top. Moreover, for statistical purposes, data referring to this condition were sign-inverted.

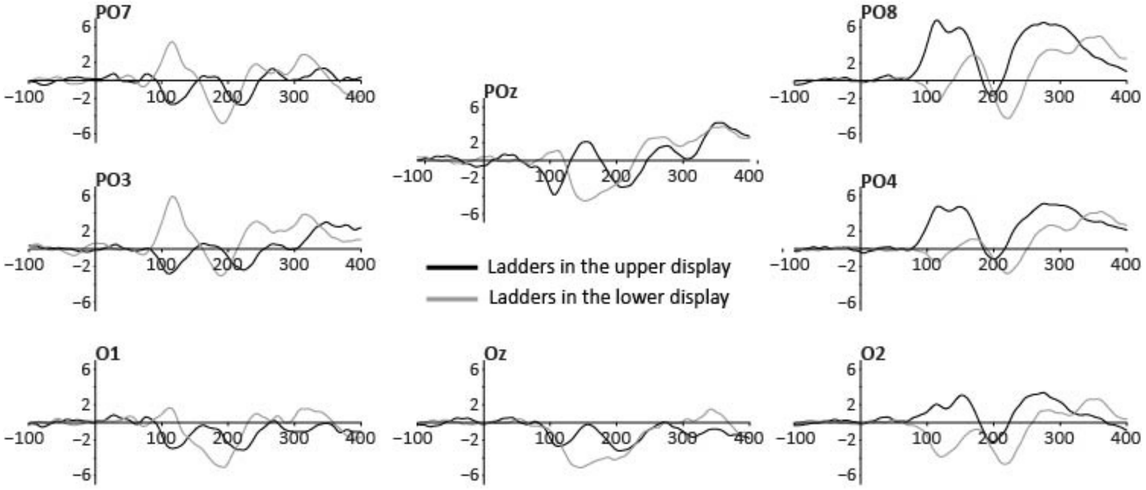


Figure 2. Inversion effect on early and late components. VEPs in the ladder static condition are showed separately for eight occipital and parieto-occipital channels. Black lines are VEPs of ladders presented in the upper-left display. Grey lines are correlated of ladders in the bottom-right display.

3. Results

3.1 Psychophysics

Figure 2 shows averaged data from 14 observers. We plotted the proportion of correct answers separately for each condition. Results show a (negligible) better performance in the dynamic condition (grey and black framed bars) for both snakes and ladders.

More evident is instead that snakes (grey bars) are always more easily detected than ladders (black bars).

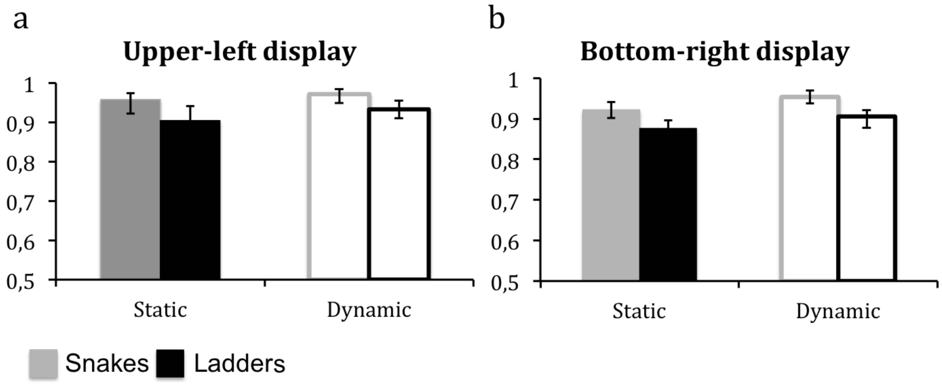


Figure 3. Behavioural data from the experiment. Averaged proportion of correct answers are shown separately for each condition. The first panel refers to the "target in the upper-left display" condition, whereas the second panel shows data from the "target in the lower-right display". Grey bars indicate snake contours, black bars ladder contours. Solid bars refer to the static condition, whereas empty bars show results in the dynamic condition. Error bars display the standard errors.

These observations were confirmed by a repeated-measure ANOVA with Motion (static vs. dynamic), Stimulus Type (snakes vs. ladders) and Position (up-left vs. down-right) as main factors. We found indeed a significant effect of Stimulus Type ($F_{(1,13)}=6.614$, $p=0.028$) and Motion ($F_{(1,13)}=5.609$, $p=0.028$). The position of the stimulus display significantly affected the performance ($F_{(1,13)}=9.859$, $p=0.014$), with worse detection when the target was displayed in the bottom-right quadrant. I will come back later on this finding. None of the interactions were significant. These findings are in agreement with previous literature on static and moving contours (Bex et al., 2001; Dakin & Baruch, 2009; Field et al., 1993; Ledgeway & Hess, 2006; Ledgeway et al., 2005; May & Hess, 2007a, 2007b, 2008; Nygard et al., 2009): we found a superiority-snake effect and motion exhibits a significant trend.

3.2 Visual Evoked Potentials

VEPs components are shown separately for the static (Figure 4) and dynamic (Figure 5) condition. Moreover, VEPs in the upper-left and lower-right conditions are showed on separate plots. Three main effects emerge: the first clearly is the general increment in negativity in correspondence of 100-200 ms when the target is displayed (black line) that is consistent with previous studies (Bach & Meigen, 1992, 1997; Casco et al., 2009; Machilsen et al., 2011; Mathes et al., 2006; Shpaner et al., 2013). The second is a general trend towards negativity found with moving stimuli (Figure 5). The third is even more interesting, although less evident: it is a different modulation of C1 amplitude in response to snakes and ladders in the static condition (Figure 4). If we take a look at the difference between contour+background (black line) and background (grey line), we can see that it shows a greater C1 positivity for snakes, when presented in the upper-left display (Figure 4, top row).

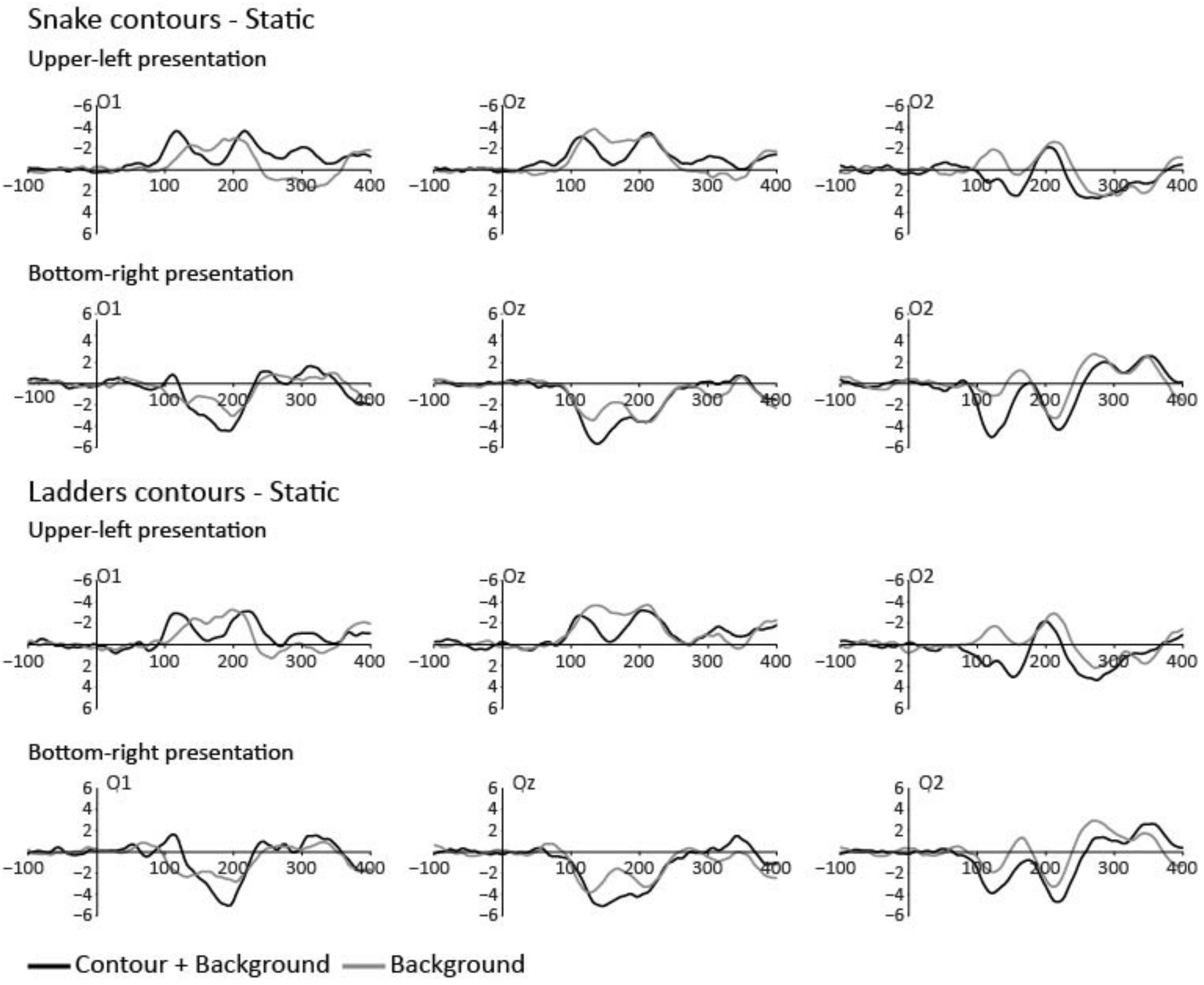


Figure 4. Electrophysiological results in the static condition, VEPs. Comparisons between VEPs in the contour+background (black line) and background-only (grey line) conditions. Grand average event-related potentials (μV) are displayed as a function of latency separately for 3 channels (column) and presentation displays (upper-left vs. bottom right; rows). The upper panel shows data for snake contours, the lower panel shows data for ladders.

However, when stimuli are presented in lower-right position (Figure 4, second row), positivity is no longer discernible and a more pronounced C1 negativity for ladders appears (Figure 4, fourth row). Following Casco et al. (2009), positivity on C1 for collinear stimuli might indicate an early grouping mechanism. The fact that ladders do not exhibit this effect is also in line with their results. The negative modulation visible with ladders, instead, could reflect other processes (e.g. background influences) in contour detection. I will come back to this point in the Discussion section.

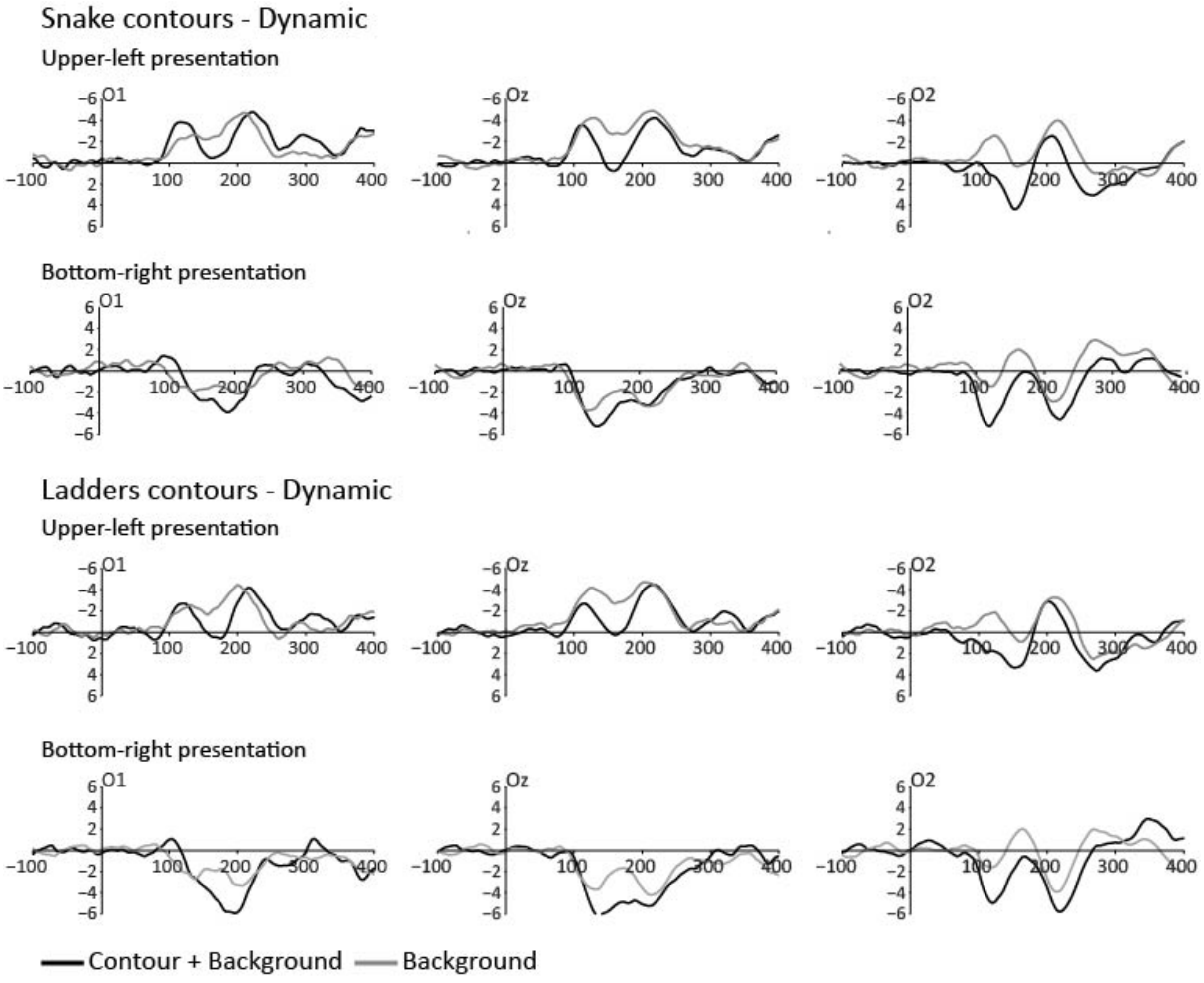


Figure 5. Electrophysiological results in the dynamic condition, VEPs. Comparisons between VEPs in the contour+background (black line) and background-only (grey line) conditions. Grand average event-related potentials (μV) are displayed as a function of latency separately for 3 channels and presentation displays are organised as in Figure 4.

To test for specific effects of motion and contour, I compared D-waves in these conditions (Figure 4) for both amplitude and latency domains. Since contour type seems to exert a different modulation on C1 according to the presentation site and since behavioural data pointed out a detrimental effect of the bottom-right presentation, analysis were conducted in isolation for upper-left and bottom-right stimuli. This allowed both testing for an eventual contour effect and avoiding confounding effects due to the anisotropy.

3.2.1 Comparison between static and dynamic integration of snakes and ladders: amplitudes

To summarise, comparisons between amplitudes in the static and dynamic condition revealed that motion generally enhanced negativity when stimuli were presented in the upper-left part of the visual field. This effect was not visible for stimuli presented in the bottom-right display. Moreover, mean amplitude of the C1 for static snakes showed a trend (even if not significant) towards positivity. The negative modulation of the C1 by ladders presented in the lower visual field was also partially confirmed.

Two separated four-way repeated-measure ANOVA on D-waves with Motion (static vs. dynamic), Contour Type (snake vs. ladder) and Component (C1 vs. N1) as factors, were conducted for the “upper-left” and “bottom-right ”conditions, separately for each channel.

In the upper-left presentation condition, I trivially found a main effect of the Component on the three channels ($F_{(1,13)}$=13.370, p=0.003 in O1; $F_{(1,13)}$=31.874, p<0.0001 in O2; $F_{(1,13)}$=23.645, p<0.0001 in Oz). In O1 also Motion and Contour Type had a main effect, with moving stimuli generally more negative than static ones ($F_{(1,13)}$=7.931, p=0.015) and ladders generally more negative than snakes ($F_{(1,13)}$=6.208, p=0.027). Similarly, in O2 ($F_{(1,3)}$=6.436, p=0.025) and Oz ($F_{(1,3)}$=11.709, p=0.005) moving stimuli were generally associated to higher negativity.

As a general rule, three-way interactions at these sites did not reach significance, thus not demonstrating what I underlined in VEPs, i.e. that static snakes have more positive amplitude in C1 than ladders. Nonetheless, mean amplitudes of D-waves showed opposite trends for snakes and ladders, with static snakes having mean C1 amplitude of -0,43 μV in O1, 0.32 μV in O2 and 0.52 μV in Oz, in contrast with ladders having of -1.49 μV in O1, -1.39 μV in O2 and -0.86 μV in Oz (Table 1). This effect, even though not statistically significant, is in agreement with (Casco et al., 2009).

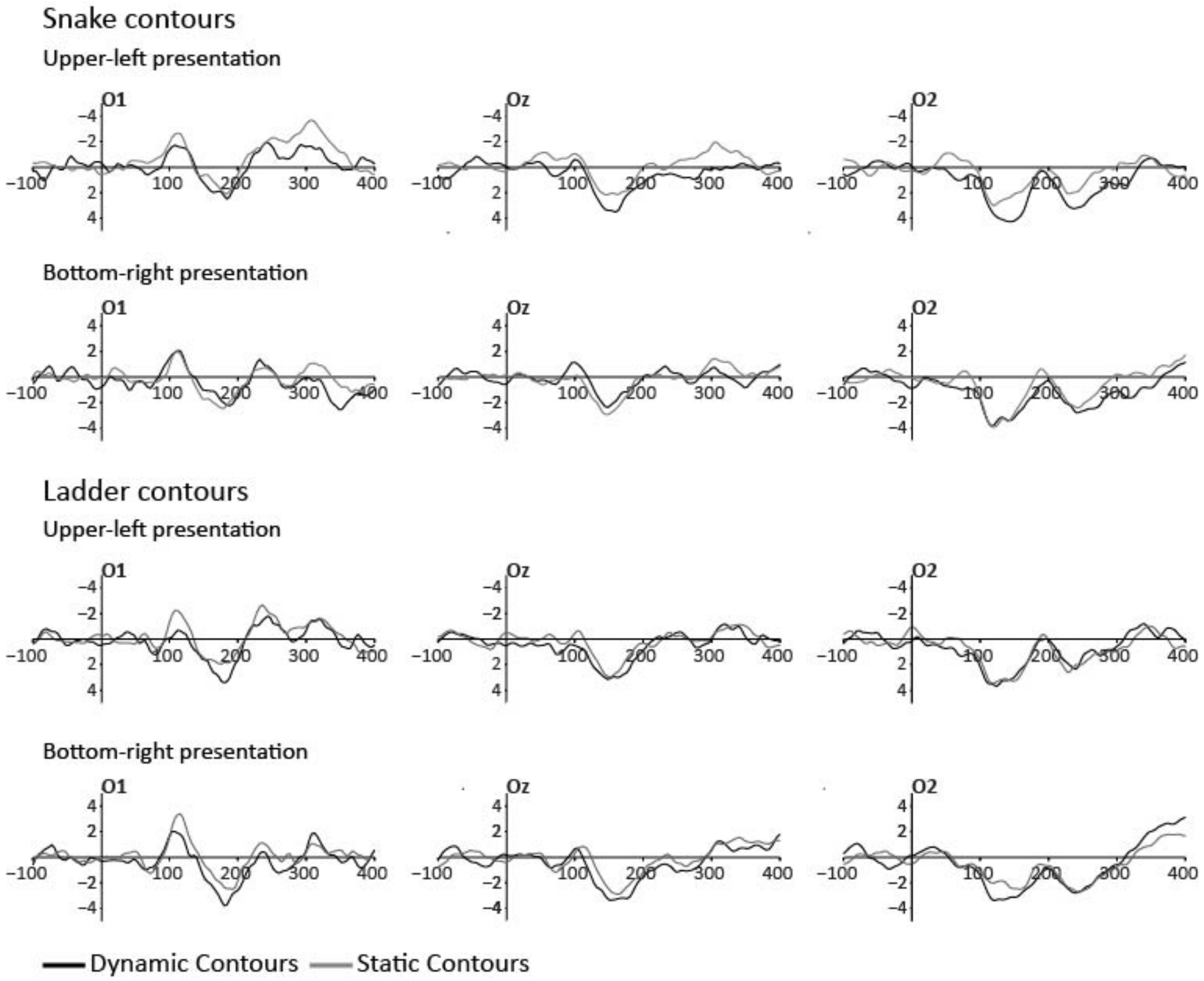
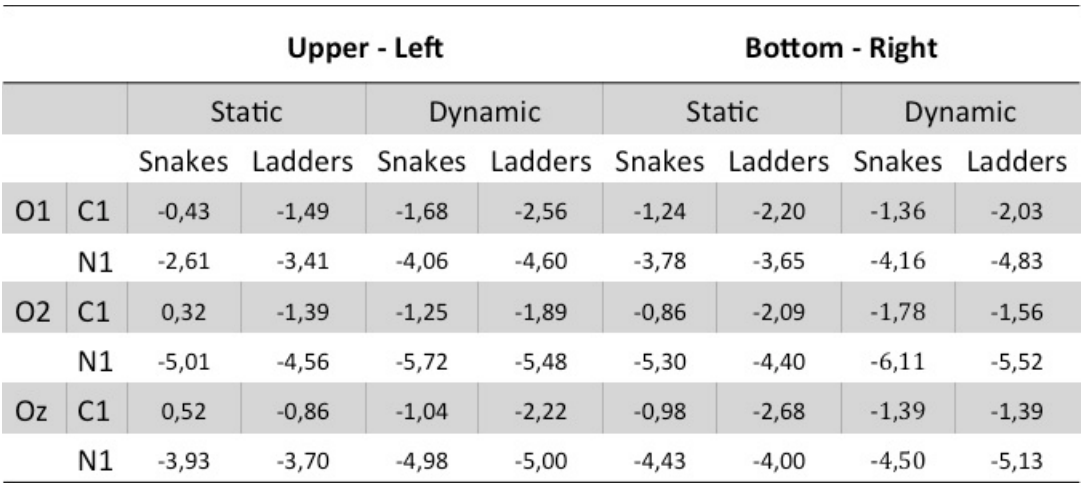


Figure 6. Electrophysiological results, D-waves. Grand average event-related potentials (μV) are displayed as a function of latency separately for 3 channels. (a) Snakes in the static (grey line) and dynamic (black line) conditions are compared separately for upper-left and bottom-right presentation. (b) Ladders in the static (grey line) and dynamic (black line) separately in the upper-left and bottom-right presentation.

A different pattern of results was instead found with stimuli in the lower-right visual field.

Again, we found a trivial main effect of the Component on the three channels ($F_{(1,13)}$=43.239, p<0.0001 in O1; $F_{(1,13)}$=34.9812, p<0.0001 in O2; $F_{(1,13)}$=23.508, p<0.0001 in Oz). However, I also found a significant interaction between Motion, Component and Contour Type in Oz ($F_{(1,13)}$=8.628, p=0.012). However, Bonferroni-corrected t-tests only reported that static ladders tended (p=0.051) to be more negative than static snakes, as showed by VEPs. None of the other interactions reached significance.

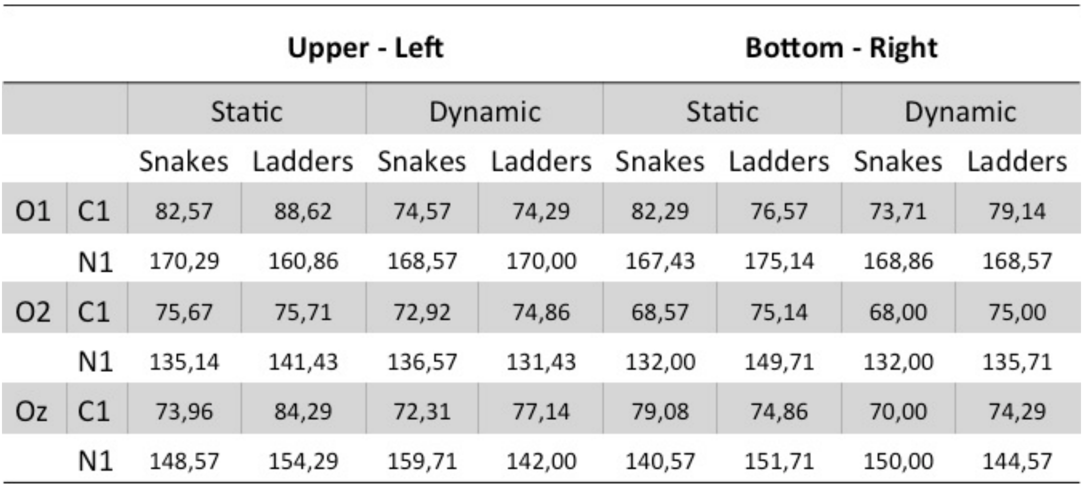


		Upper - Left				Bottom - Right			
		Static		Dynamic		Static		Dynamic	
		Snakes	Ladders	Snakes	Ladders	Snakes	Ladders	Snakes	Ladders
O1	C1	-0,43	-1,49	-1,68	-2,56	-1,24	-2,20	-1,36	-2,03
	N1	-2,61	-3,41	-4,06	-4,60	-3,78	-3,65	-4,16	-4,83
O2	C1	0,32	-1,39	-1,25	-1,89	-0,86	-2,09	-1,78	-1,56
	N1	-5,01	-4,56	-5,72	-5,48	-5,30	-4,40	-6,11	-5,52
Oz	C1	0,52	-0,86	-1,04	-2,22	-0,98	-2,68	-1,39	-1,39
	N1	-3,93	-3,70	-4,98	-5,00	-4,43	-4,00	-4,50	-5,13

Table 1. D-waves averaged amplitudes (ms) for static Snakes, static Ladders, separately for upper-left and bottom-right presentation (left and right panel, respectively).

3.2.2 Comparison between static and dynamic integration of snakes and ladders: latencies

Effects of motion and contour type on latencies were generally less clear. Again, two separated three-way repeated-measure ANOVA on D-waves with Motion (static vs. dynamic), Contour Type (snake vs. ladder) and Component (C1 vs. N1) as factors, were conducted for the “upper-left” and “bottom-right ”conditions and separately for O1, O2 and Oz.



		Upper - Left				Bottom - Right			
		Static		Dynamic		Static		Dynamic	
		Snakes	Ladders	Snakes	Ladders	Snakes	Ladders	Snakes	Ladders
O1	C1	82,57	88,62	74,57	74,29	82,29	76,57	73,71	79,14
	N1	170,29	160,86	168,57	170,00	167,43	175,14	168,86	168,57
O2	C1	75,67	75,71	72,92	74,86	68,57	75,14	68,00	75,00
	N1	135,14	141,43	136,57	131,43	132,00	149,71	132,00	135,71
Oz	C1	73,96	84,29	72,31	77,14	79,08	74,86	70,00	74,29
	N1	148,57	154,29	159,71	142,00	140,57	151,71	150,00	144,57

Table 2. D-waves averaged latencies (ms) for static Snakes, static Ladders, separately for upper-left and bottom-right presentation (left and right panel, respectively).

In both the display conditions, as with amplitudes, Component always resulted in an obvious main effect (Upper-left: $F_{(1,13)}=652.150$, $p<0.0001$ in O1; $F_{(1,13)}=293.222$, $p<0.0001$ in O2; $F_{(1,13)}=671.522$, $p<0.0001$ in Oz. Bottom-right: $F_{(1,13)}=376.217$, $p<0.0001$ in O1; $F_{(1,13)}=228.635$, $p<0.0001$ in O2; $F_{(1,13)}=276.313$, $p<0.0001$ in Oz).

In the upper-left condition, in O1 also the interaction between Motion and Component reached significance ($F_{(1,13)}=7.375$, $p=0.018$): this indicates that dynamic stimuli (either snakes or ladders) had shorter C1 latencies ($p=0.003$). In Oz we found significant interactions between Motion and Contour ($F_{(1,13)}=5.385$, $p=0.037$) and between Component and Contour ($F_{(1,13)}=5.596$, $p=0.035$): Bonferroni corrected pairwise comparisons revealed that motion generally speeded up ladder latencies ($p=0.037$). Finally, snakes resulted faster than ladders in the C1 recorded in this site ($p=0.021$), regardless of motion.

With stimuli in the lower-right quadrant, we found a main effect of Contour Type in O2 ($F_{(1,13)}=4.702$, $p=0.049$), showing that snakes were generally faster than ladders, regardless of motion. In Oz a three-way interactions (Motion vs. Component vs. Contour) reached significance ($F_{(1,13)}=$, $p=0.028$), only indicating that the two components emerged at significantly different times after stimulus onset ($p<0.0001$ in all the conditions). No other significant interactions were found.

4. Discussion

To summarise, I tested observers in a contour-detection task with either static or moving contours. Both behavioural and electrophysiological measures were compared in the two conditions. Better performance was generally found in correspondence of dynamic and snake contours, in line with previous studies (Bex et al., 2001; Dakin & Baruch, 2009; Field et al., 1993; Ledgeway & Hess, 2006; Ledgeway et al., 2005; May & Hess, 2007a, 2007b, 2008; Nygard et al., 2009). Moreover, I also found an effect of display position, with lower accuracy usually associated with stimuli presented in the lower visual field. Electrophysiological results especially pointed out a negative deflection of the D-waves in the dynamic condition. Such negativity was constant when stimuli were presented in the upper-left visual field, regardless of contour type. Latencies were in some conditions longer for ladders, in agreement with data by (Tanskanen et al., 2008), who found slower contour-selective response for radial than for collinear stimuli. The difference in contour types and motion and static conditions are discussed below.

4.1. Different visual evoked correlates for snakes and ladders

This study investigated the electrophysiological correlates of snakes and ladders when contours were static or motion-defined. The activity generated by the contour was isolated with a subtractive method, so that local signals (amount of elements, range of local orientations and, consequently, of direction of motion) were removed. Thus, any difference in static and dynamic conditions reflects differences in *contour integration*. As already pointed out by Casco et al. (2009), however, two processes need to be taken into account: the first is grouping of elements sharing some feature. The second is segmentation of the group from the background. In their study, grouping was reported to be early and to increase positivity of C1 for collinear stimuli, whereas segmentation was repeatedly mirrored by an enhancement of N1 amplitude (Bach & Meigen, 1992, 1997; Caputo & Casco, 1999; Casco et al., 2009; Machilsen et al., 2011; Mathes et al., 2006; Shpaner et al., 2013). In this preliminary work I have found that grouping in snakes modulated both amplitude (in O1, O2 and Oz) and latency (in Oz) of the earliest component (C1). However, the effect was not significant, possibly because our stimuli differed in many respects from those used by Casco et al (contour length, eccentricity, background type; I will come back to this). Surprisingly also ladders modulated C1 amplitude in a way opposite to snakes: in the static condition, they were correlated to a more negative C1 than snakes.

Before discussing this "inverse modulation" effect on the C1, it is worth mentioning that previous studies failed at finding any variation of the C1 amplitude due to contour features. A couple differences can account for this discrepancy: in some cases, C1 modulation wasn't analysed (Machilsen et al., 2011; Mathes et al., 2006). When instead focus was specifically on early components as in (Shpaner et al., 2013), ladders were not used. Another difference between this study and Shpaner and col's could rely in the fact that they used curved contours, whereas I used rather simple straight contours, specifically designed to promote optimal activation of static and dynamic association field, in early visual areas (see Chapter 3; (W. Li & Gilbert, 2002; W. Li et al., 2006)).

To better understand the “inverse modulation” effect on the C1, one need to step back to “association field” models of contour integration. One influential model is the one proposed by Yen and Finkel (1998) (see also Chapter 2). The authors explicitly included two types of excitatory, long-range, horizontal connections: the “co-axial” connections link collinear receptive field tuned for smooth curves; the “trans-axial” ones connect receptive fields that have the same preferred orientation, but are orthogonal to those curves. In other words, connections for both snakes (via ‘‘co-axial’’) and ladders (via ‘‘trans-axial’’) are provided. An important feature of this model is also the presence of inhibitory connections between elements belonging the contour and elements in the background. Inhibition and excitation compete in a “winner-takes-all” process, in which the algorithm compares activation between configurations and switches off the less-active path. Another important feature is that trans-axial connections are weaker than co-axial ones. Thus, contrarily to the seminal model by Field et al. (1993) that only suggests the presence of co-axial links, this one can account for the integration of ladders and can at the same time explain the snake “superiority” effect.

These two features may justify my findings. When a ladder is presented in a detection task, the visual system is required, on one hand, to enhance inhibition towards distracters and activation between elements in the contour configuration. On the other hand, weak trans-axial links make ladders more vulnerable to distracters. Thus, distracters inhibition needs to be incremented. The detrimental contextual effect exerted by the background has recently been found by a number of studies (Dakin & Baruch, 2009; May & Hess, 2007b, 2008; Robol et al., 2012; Ungerleider & Pasternak, 2004) (Chapter 3), with authors concluding that distracters induce crowding of the contour. Thus, suppression represents a good candidate to explain the early negativity associated to static ladders. Moreover, it also underlines the importance of local orientation and contextual effects in contour integration tasks.

Other cues indicate that this interpretation may hold: first, it is known that crowding in general is bigger in the periphery of the visual field (Levi, 2008) and our stimuli were presented outside the foveal vision. Moreover, iso-oriented stimuli induce pop-out of the target (Knierim & van Essen, 1992), therefore the iso-oriented background used by Casco et al. (2009) could have reduced this crowding effect. Machilsen et al. (2011) specifically tested modulation in ERPs response to contour stimuli in presence of either iso- or randomly-oriented backgrounds. They found background to influence N1, with iso-orientation inducing a more negative peak, and P1. However, they did not assess modulation of earlier component.

4.2. Electrophysiological correlates of motion facilitation

One main result from this study is that motion enhances negativity of the contour-related components, both on early (C1) and later (N1) time window.

The negative deflection on N1 could indicate that motion generally helps segmentation, making stimuli more salient. This is again in agreement with the idea that an enhancement in N1 negativity mirrors pop-out (Herrmann & Bosch, 2001; Hillyard & Anllo-Vento, 1998; Hopf, Boelmans, Schoenfeld, Luck, & Heinze, 2004; S. J. Luck, Heinze, Mangun, & Hillyard, 1990; Mathes et al., 2006; Senkowski, Rottger, Grimm, Foxe, & Herrmann, 2005). Mathes et al. (2006), for example, found more negative N1 in a contour-detection task when the target was easy to detect, rather then when task demands were high. According to the authors, an N1 enhancement might indicate fast segmentation processes and this would explain shorter latencies in the snake condition as well. It has also been showed that a Kanizsa figure that automatically attracts attention when embedded in a field of distracters elicits a negative modulation in a time window corresponding to N1, further indicating a relation between pop-out and N1 (Senkowski et al., 2005). Also Casco et al. (2005) found larger N1 amplitude in a texture segmentation task for attended locations and interpreted it as enhanced salience in this condition.

Consistently, I generally found shorter latencies in correspondence of moving ladders, this being a correlate of motion facilitation observed with these stimuli. What is surprising, however, is that dynamic stimuli also modulate C1 amplitude. This early effect further reflects what I found with static stimuli, i.e. that the increment in negativity could somehow indicate a release from distracters inhibition on target. This then results in more efficient grouping and segmentation of the contour and, thus, in a greater amount of correct answers in this condition. As a consequence of more efficient detection, also latencies are reasonably reduced: indeed, in O1 dynamic contours had shorter C1 latencies than static ones. This provides an answer to the question “how it is that motion facilitates contour integration”.

A correlate of this interpretation is that motion intervenes early in visual processing, i.e. before segmentation, at the level of lateral interactions amongst receptive fields in V1 (Fitzpatrick, 1996; Gilbert & Wiesel, 1989; Kapadia et al., 1995; Kastner et al., 1999).
In Chapter 4 I discussed the possibility that a cooperative mechanism, pooling together signals sharing the same direction, accounts for the integration of dynamic contours.
I suggested that this mechanism acts at a local and early level, prior to the integration of local orientation. The findings in this study could provide electrophysiological correlates of cooperativity. The early processing might reflect the modulation I found as soon as of C1 amplitude. Its relation to local orientations might further sustain the hypothesis that motion facilitation consists in reducing crowding possibly by strengthening (in this case) local target signals having the same direction. The neural correlate of this strengthening might be the negative deflection we found with dynamic contours.

It should be remarked, however, that these data only refer to three occipital sites and to a short-latency after stimulus onset. The focus was indeed to investigate early modulation as correlates of motion facilitation. Nonetheless, an effect of motion is also visible on parietal sites (see figure 4). Finally, since motion onset elicits a later negative component (N2, see (Heinrich, 2007) for a review), future analysis should include also later components, to test also for the presence of recurrent feedback modulation influencing early processing.

4.3. Is lateralisation an accountable explanation for differences in the upper-left and lower-right presentation sites?

Behavioural data showed, unfortunately, a detrimental effect of the bottom-right presentation. Moreover, I didn't find modulation of VEPs by motion in this presentation condition. What emerges, then, is an advantage for stimuli presented in the visual left field that therefore project to the right visual hemisphere. A possible explanation comes from studies on brain-damaged patients: patients with left hemisphere lesion performed as controls in a grouping task, whereas right-damaged patients failed at this task (Vancleef, Wagemans, & Humphreys, 2013), clearly showing a major recruitment of the right hemisphere in grouping tasks. The picture is complicated, however, by the fact that right-damaged patients only showed a small decrement in contour detection performance (probably because of recurrent connections from higher visual areas). Nonetheless, grouping represents a main feature involved in this study, especially given the particular simplicity of straight contours.

Another asymmetry, however, might be responsible for these results. A general advantage of lower visual field is indeed well-known ((He, Cavanagh, & Intriligator, 1996; Levine & McAnany, 2005; McAnany & Levine, 2007; Skrandies, 1987); see (Christman & Niebauer, 1997) for a review on upper-lower asymmetries). It has been suggested that the asymmetry might lie in projections from different part of the visual field to magno- and parvo-cellular systems (Pitts et al., 2011). However, it has also been documented a different effect of attentional modulation on C1 caused by asymmetries in the (upper vs. lower) visual field (Rauss, Pourtois, Vuilleumier, & Schwartz, 2009). Finally, Thomas and Elias (2011) found a leftward bias (i.e. a general attentional bias towards the leftward visual field (Jewell & McCourt, 2000)) to be higher for upper stimuli at short presentation, while prolonged presentation moved the bias to the lower field. Thus, the short stimulus presentation in this experiment (200 ms) might have biased performance towards the upper field. However, my findings do not provide any specific test to disentangle the asymmetry. The lateralisation issue is still an open question and further investigation is required.

4.4. Conclusions

To summarise, preliminary data from this experiment show that snakes and ladders exhibit different polarity modulation on early visual components. Ladders are associated to more negativity, whereas amplitudes linked to snake presentation show a trend towards larger positivity. I suggest that while the modulation of early components with snakes may reflect facilitation amongst contour elements (i.e. by association field), with ladders it may reflect reduced inhibition of background elements (i.e. reduced crowding).

Furthermore, motion facilitation is associated to a general increase in negativity (at least on three occipital channels). Presumably motion reduces time necessary for detection as well. Moreover, negativity can appear at different times after stimulus onset (indeed it modulates both the analysed components). The C1 effect suggests higher efficiency in reducing crowding, whereas N1 modulation is indicative of higher efficiency in segmentation. The former effect is in agreement with the idea that a local, cooperative, process is responsible for moving contour integration. However, further investigations are needed to assess whether this early modulation is actually generated by low-level areas or is inherited by higher-areas via recurrent connections.

General discussion

Perceptual grouping and contour integration have been widely tested. Nonetheless, there are still open questions. One of them refers to the mechanism responsible of integrating static disconnected elements into contours, that should be able to explain contextual effects as well as differences in integration for differently-defined contours. The second is related to the interplay that occurs between static and dynamic local signals.

From studies one and three it emerges that a model based on connections between similarly oriented and adjacent receptive fields, rather than one based on second-order filtering, can account for a series of results. Specifically, a model based on the balance between excitatory and inhibitory lateral interactions, as the extension of the association field proposed by Yen and Finkel (1998), could account for the detection of both contours. This model contains both excitatory and inhibitory connections between receptive fields having respectively same and different orientations; moreover, connections between parallel oriented receptive fields are weaker than connections between collinear receptive fields.

First of all, it could explain the detrimental effect of background modulation (at low spatial frequency) and of phase randomization found in study one. I manipulated the arrangement of the background in a contour-detection task, to avoid grouping by position. Detection with both snakes and ladders was tested. Results showed that, provided sufficient orientation information, position was no longer a cue required by the system to perform detection. In contrast, at low spatial frequency, performance was strongly reduced when position was missing. Furthermore, regardless of the spatial organization of the background, snakes were more easily detected when all the elements were in phase. Ladders, instead, were slightly but consistently improved by phase randomization. A second-order filter (following rectification of first-order filters) is insensitive to local properties and contextual influences and was therefore rejected.

In study three I recorded visual evoked responses (VEPs) to snakes and ladders. I found a different correlate of grouping in the two cases: snakes were characterised by an early positive modulation, suggesting that grouping occurs in this condition (Casco et al., 2009). The same time-window was instead characterised by negativity for ladders. According to the model by Yen and Finkel, I suggested that background exerts more inhibition on ladder contours, integrated by weak trans-axial connections, than on snakes. Thus, the system is required to reduce suppression from the background activation and this represents a good candidate to explain the early negativity associated to static ladders. This is in line with a series of studies showing impairment on contour detection by background interference (Dakin & Baruch, 2009; May & Hess, 2007b, 2008; Robol et al., 2012; Ungerleider & Pasternak, 2004), with authors concluding that distracters induce crowding of the contour. Interestingly, the facilitation effect by phase randomization found with ladders (Study one) could be also explained as a release from crowding.

Study two focused, instead, on the interactions between static and dynamic local signals. My results showed, for the first time, an active influence of motion on static signals. Experiment three in particular showed that (alternate) motion could perturb a static attribute of the static configuration (i.e. positions of the local elements), with observers less able to detect contours in this rather than in the static condition. Findings from experiment 3 (i.e. alternating local direction produces perturbation in the perceived positions of the elements along the contour) therefore indicate that motion is extracted *prior to* static configuration. The consequence is a non-optimal activation of the *static* association field, given that the orientation-position constrained is no longer respected. Consistently, motion could enhance links within those elements coherently moving, helping the static association field and thus explaining the facilitation in detection that usually occurs with motion. Thus, according to my data, motion extraction seems to intervene before and modulate static association field operations. These findings were also mirrored by the negative deflection found with dynamic contours in study three. Negative amplitude on early time window for ladders indicated more inhibition of the surround. More negativity on the segmentation component could be interpreted as an increment of contour salience caused by the enhancement of connections within coherently moving elements.

I also collected evidence that a cooperative motion mechanism integrates coherently moving local motion signals (Vreven & Verghese, 2002). Study two showed that an increment in the amount of dynamic signals leads to an increment of detection of suprathreshold stimuli. Moreover, the importance of local direction suggested by this and previous studies (Bex et al., 2001; R. F. Hess & Ledgeway, 2003; Ledgeway & Hess, 2002; Vreven & Verghese, 2002) further support this hypothesis.

7. References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A, 2*(2), 284-299.

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature, 300*(5892), 523-525.

Allman, J., Miezin, F., & McGuinness, E. (1985). Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception, 14*(2), 105-126.

Amano, K., Edwards, M., Badcock, D. R., & Nishida, S. (2009). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *J Vis, 9*(3), 4 1-25. doi: 10.1167/9.3.4

Amano, K., Goda, N., Nishida, S., Ejima, Y., Takeda, T., & Ohtani, Y. (2006). Estimation of the timing of human visual perception from magnetoencephalography. *J Neurosci, 26*(15), 3981-3991. doi: 10.1523/JNEUROSCI.4343-05.2006

Amano, K., Kuriki, I., & Takeda, T. (2005). Direction-specific adaptation of magnetic responses to motion onset. *Vision Res, 45*(19), 2533-2548. doi: 10.1016/j.visres.2005.02.024

Anderson, S. J., & Burr, D. C. (1987). Receptive field size of human motion detection units. *Vision Res, 27*(4), 621-635.

Anderson, S. J., Holliday, I. E., Singh, K. D., & Harding, G. F. (1996). Localization and functional analysis of human cortical area V5 using magneto-encephalography. *Proc Biol Sci, 263*(1369), 423-431. doi: 10.1098/rspb.1996.0064

Aspell, J. E., Tanskanen, T., & Hurlbert, A. C. (2005). Neuromagnetic correlates of visual motion coherence. *Eur J Neurosci, 22*(11), 2937-2945. doi: 10.1111/j.1460-9568.2005.04473.x

Attneave, F. (1954). Some informational aspects of visual perception. *Psychol Rev, 61*(3), 183-193.

Bach, M., & Meigen, T. (1992). Electrophysiological correlates of texture segregation in the human visual evoked potential. *Vision Res, 32*(3), 417-424.

Bach, M., & Meigen, T. (1997). Similar electrophysiological correlates of texture segregation induced by luminance, orientation, motion and stereo. *Vision Res, 37*(11), 1409-1414.

Bach, M., & Ullrich, D. (1997). Contrast dependency of motion-onset and pattern-reversal VEPs: interaction of stimulus type, recording site and response component. *Vision Res, 37*(13), 1845-1849.

Banton, T., & Bertenthal, B. I. (1996). Infants' sensitivity to uniform motion. *Vision Res, 36*(11), 1633-1640.

Bell, J., Gheorghiu, E., Hess, R. F., & Kingdom, F. A. (2011). Global shape processing involves a hierarchy of integration stages. *Vision Res, 51*(15), 1760-1766. doi: 10.1016/j.visres.2011.06.003

Berkson, J. (1953). A statistically precise and relatively simple method of estimating the bio-assay with quantal response, based on the logistic function. *Journal of the American Statistical Association, 48*(263), 565-599.

Bex, P. J., & Dakin, S. C. (2002). Comparison of the spatial-frequency selectivity of local and global motion detectors. *J Opt Soc Am A Opt Image Sci Vis, 19*(4), 670-677.

Bex, P. J., Simmers, A. J., & Dakin, S. C. (2001). Snakes and ladders: the role of temporal modulation in visual contour integration. *Vision Res, 41*(27), 3775-3782.

Bex, P. J., Simmers, A. J., & Dakin, S. C. (2003). Grouping local directional signals into moving contours. *Vision Res, 43*(20), 2141-2153.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychol Rev, 94*(2), 115-147.

Bonneh, Y., & Sagi, D. (1998). Effects of spatial configuration on contrast detection. *Vision Res, 38*(22), 3541-3553.

Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annu Rev Neurosci, 28*, 157-189. doi: 10.1146/annurev.neuro.26.041002.131052

Born, R. T., & Tootell, R. B. (1992). Segregation of global and local motion processing in primate middle temporal visual area. *Nature, 357*(6378), 497-499. doi: 10.1038/357497a0

Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J Neurosci, 17*(6), 2112-2127.

Bowns, L., & Alais, D. (2006). Large shifts in perceived motion direction reveal multiple global motion solutions. *Vision Res, 46*(8-9), 1170-1177. doi: 10.1016/j.visres.2005.08.029

Bozzi, P. (1969). Direzionalità ed organizzazione interna della figura. [Directionality and internal organization of a figure]. *Atti dell'Accademia patavina di Scienze, Lettere ed Arti, 81*, 137-170.

Bozzi, P. (1989). *Fenomenologia sperimentale*. Bologna: Il Mulino.

Braddick, O. J., O'Brien, J. M., Wattam-Bell, J., Atkinson, J., & Turner, R. (2000). Form and motion coherence activate independent, but not dorsal/ventral segregated, networks in the human brain. *Curr Biol, 10*(12), 731-734.

Bradley, D. C., & Andersen, R. A. (1998). Center-surround antagonism based on disparity in primate area MT. *J Neurosci, 18*(18), 7552-7565.

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

Bruce, V. (2003). *Visual perception: Physiology, psychology, and ecology.* : Psychology Press.

Bundo, M., Kaneoke, Y., Inao, S., Yoshida, J., Nakamura, A., & Kakigi, R. (2000). Human visual motion areas determined individually by magnetoencephalography and 3D magnetic resonance imaging. *Hum Brain Mapp, 11*(1), 33-45.

Caputo, G., & Casco, C. (1999). A visual evoked potential correlate of global figure-ground segmentation. *Vision Res, 39*(9), 1597-1610.

Casco, C., Campana, G., Grieco, A., & Fuggetta, G. (2004). Perceptual learning modulates electrophysiological and psychophysical response to visual texture segmentation in humans. *Neurosci Lett, 371*(1), 18-23. doi: 10.1016/j.neulet.2004.08.005

Casco, C., Campana, G., Han, S., & Guzzon, D. (2009). Psychophysical and electrophysiological evidence of independent facilitation by collinearity and similarity in texture grouping and segmentation. *Vision Res, 49*(6), 583-593. doi: 10.1016/j.visres.2009.02.004

Casco, C., Caputo, G., & Grieco, A. (2001). Discrimination of an orientation difference in dynamic textures. *Vision Res, 41*(3), 275-284.

Casco, C., Grieco, A., Campana, G., Corvino, M. P., & Caputo, G. (2005). Attention modulates psychophysical and electrophysiological response to visual texture segmentation in humans. *Vision Res, 45*(18), 2384-2396. doi: 10.1016/j.visres.2005.02.022

Casco, C., Grieco, A., Giora, E., & Martinelli, M. (2006). Saliency from orthogonal velocity component in texture segregation. *Vision Res, 46*(6-7), 1091-1098. doi: 10.1016/j.visres.2005.09.032

Chakravarthi, R., & Pelli, D. G. (2011). The same binding in contour integration and crowding. *J Vis, 11*(8). doi: 10.1167/11.8.10

Champion, R. A., Hammett, S. T., & Thompson, P. G. (2007). Perceived direction of plaid motion is not predicted by component speeds. *Vision Res, 47*(3), 375-383. doi: 10.1016/j.visres.2006.10.017

Chang, J. J., & Julesz, B. (1984). Cooperative phenomena in apparent movement perception of random-dot cinematograms. *Vision Res, 24*(12), 1781-1788.

Chatterjee, S., & Price, B. (1977). *Regression analysis by example*. New York: John Wiley.

Chen, C. C., & Tyler, C. W. (1999). Spatial pattern summation is phase-insensitive in the fovea but not in the periphery. *Spat Vis, 12*(3), 267-285.

Chicherov, V., Plomp, G., & Herzog, M. H. (2011). *The time course of perceptual grouping: a high density ERP study.* Paper presented at the European Conference of Visual Perception, Tolouse.

Christman, S. D., & Niebauer, C. L. (1997). The relation between left-right and upper-lower visual field asymmetries: or: What goes up goes right, while what's left lays low. *Advances in Psychology, 123*, 263-296.

Chung, S. T., Patel, S. S., Bedell, H. E., & Yilmaz, O. (2007). Spatial and temporal properties of the illusory motion-induced position shift for drifting stimuli. *Vision Res, 47*(2), 231-243. doi: 10.1016/j.visres.2006.10.008

Clark, V. P., Fan, S., & Hillyard, S. A. (1994). Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping, 2*(3), 170-187.

Culham, J., He, S., Dukelow, S., & Verstraten, F. A. (2001). Visual motion and the human brain: what has neuroimaging told us? *Acta Psychol (Amst), 107*(1-3), 69-94.

Dakin, S. C., & Baruch, N. J. (2009). Context influences contour integration. *J Vis, 9*(2), 13 11-13. doi: 10.1167/9.2.13

Dakin, S. C., & Hess, R. F. (1998). Spatial-frequency tuning of visual contour integration. *J Opt Soc Am A Opt Image Sci Vis, 15*(6), 1486-1499.

Dakin, S. C., & Hess, R. F. (1999). Contour integration and scale combination processes in visual edge detection. *Spat Vis, 12*(3), 309-327.

Das, A., & Gilbert, C. D. (1999). Topography of contextual modulations mediated by short-range interactions in primary visual cortex. *Nature, 399*(6737), 655-661. doi: 10.1038/21371

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision Res, 31*(9), 1619-1626.

Delon-Martin, C., Gobbele, R., Buchner, H., Haug, B. A., Antal, A., Darvas, F., & Paulus, W. (2006). Temporal pattern of source activities evoked by different types of motion onset stimuli. *Neuroimage, 31*(4), 1567-1579. doi: 10.1016/j.neuroimage.2006.02.013

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods, 134*(1), 9-21. doi: 10.1016/j.jneumeth.2003.10.009

Desimone, R., & Ungerleider, L. G. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J Comp Neurol, 248*(2), 164-189. doi: 10.1002/cne.902480203

Di Russo, F., Martinez, A., & Hillyard, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cereb Cortex, 13*(5), 486-499.

Di Russo, F., Martinez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Hum Brain Mapp, 15*(2), 95-111.

Dresp, B., & Grossberg, S. (1997). Contour integration across polarities and spatial gaps: from local contrast filtering to global grouping. *Vision Res, 37*(7), 913-924.

Duffy, C. J., & Wurtz, R. H. (1991). Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. *J Neurophysiol, 65*(6), 1346-1359.

Dukelow, S. P., DeSouza, J. F., Culham, J. C., van den Berg, A. V., Menon, R. S., & Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *J Neurophysiol, 86*(4), 1991-2000.

Dupont, P., De Bruyn, B., Vandenberghe, R., Rosier, A. M., Michiels, J., Marchal, G., . . . Orban, G. A. (1997). The kinetic occipital region in human visual cortex. *Cereb Cortex, 7*(3), 283-292.

Edwards, M. (2009). Common-fate motion processing: interaction of the On and Off pathways. *Vision Res, 49*(4), 429-438. doi: 10.1016/j.visres.2008.11.010

Edwards, M., & Badcock, D. R. (1994). Global motion perception: interaction of the ON and OFF pathways. *Vision Res, 34*(21), 2849-2858.

Ellison, A., & Cowey, A. (2006). TMS can reveal contrasting functions of the dorsal and ventral visual processing streams. *Exp Brain Res, 175*(4), 618-625. doi: 10.1007/s00221-006-0582-8

Ellison, A., & Cowey, A. (2007). Time course of the involvement of the ventral and dorsal visual processing streams in a visuospatial task. *Neuropsychologia, 45*(14), 3335-3339. doi: 10.1016/j.neuropsychologia.2007.06.014

Ellison, A., & Cowey, A. (2009). Differential and co-involvement of areas of the temporal and parietal streams in visual tasks. *Neuropsychologia, 47*(6), 1609-1614. doi: 10.1016/j.neuropsychologia.2008.12.013

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. (2007). Masking disrupts reentrant processing in human visual cortex. *J Cogn Neurosci, 19*(9), 1488-1497. doi: 10.1162/jocn.2007.19.9.1488

Feldman, J., & Singh, M. (2005). Information along contours and object boundaries. *Psychol Rev, 112*(1), 243-252. doi: 10.1037/0033-295X.112.1.243

Felleman, D. J., & Van Essen, D. C. (1987). Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex. *J Neurophysiol, 57*(4), 889-920.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex, 1*(1), 1-47.

ffytche, D. H., Guy, C. N., & Zeki, S. (1995). The parallel visual motion inputs into areas V1 and V5 of human cerebral cortex. *Brain, 118 (Pt 6)*, 1375-1394.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vision Res, 33*(2), 173-193.

Field, D. J., Hayes, A., & Hess, R. F. (2000). The roles of polarity and symmetry in the perceptual grouping of contour fragments. *Spat Vis, 13*(1), 51-66.

Fitzpatrick, D. (1996). The functional organization of local circuits in visual cortex: insights from the study of tree shrew striate cortex. *Cereb Cortex, 6*(3), 329-341.

Fort, A., Besle, J., Giard, M. H., & Pernier, J. (2005). Task-dependent activation latency in human visual extrastriate cortex. *Neurosci Lett, 379*(2), 144-148. doi: 10.1016/j.neulet.2004.12.076

Foxe, J. J., Murray, M. M., & Javitt, D. C. (2005). Filling-in in schizophrenia: a high-density electrical mapping and source-analysis investigation of illusory contour processing. *Cereb Cortex, 15*(12), 1914-1927. doi: 10.1093/cercor/bhi069

Foxe, J. J., & Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans. A framework for defining "early" visual processing. *Exp Brain Res, 142*(1), 139-150. doi: 10.1007/s00221-001-0906-7

Foxe, J. J., Strugstad, E. C., Sehatpour, P., Molholm, S., Pasieka, W., Schroeder, C. E., & McCourt, M. E. (2008). Parvocellular and magnocellular contributions to the initial generators of the visual evoked potential: high-density electrical mapping of the "C1" component. *Brain Topogr, 21*(1), 11-21. doi: 10.1007/s10548-008-0063-4

Gegenfurtner, K. R., Kiper, D. C., & Fenstemaker, S. B. (1996). Processing of color, form, and motion in macaque area V2. *Vis Neurosci, 13*(1), 161-172.

Gegenfurtner, K. R., Kiper, D. C., & Levitt, J. B. (1997). Functional properties of neurons in macaque area V3. *J Neurophysiol, 77*(4), 1906-1923.

Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. *Nature, 400*(6739), 65-69. doi: 10.1038/21886

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

Gilad, A., Meirovithz, E., & Slovin, H. (2013). Population responses to contour integration: early encoding of discrete elements and late perceptual grouping. *Neuron, 78*(2), 389-402. doi: 10.1016/j.neuron.2013.02.013

Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J Neurosci, 9*(7), 2432-2442.

Gomez Gonzalez, C. M., Clark, V. P., Fan, S., Luck, S. J., & Hillyard, S. A. (1994). Sources of attention-sensitive visual event-related potentials. *Brain Topogr, 7*(1), 41-51.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci, 15*(1), 20-25.

Graham, N., Beck, J., & Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: effects of sign and amount of contrast. *Vision Res, 32*(4), 719-743.

Graham, N., & Sutter, A. (1998). Spatial summation in simple (Fourier) and complex (non-Fourier) texture channels. *Vision Res, 38*(2), 231-257.

Graham, N., Sutter, A., Venkatesan, C., & Humaran, M. (1992). Non-linear processes in perceived region segregation: orientation selectivity of complex channels. *Ophthalmic Physiol Opt, 12*(2), 142-146.

Graham, N., & Wolfson, S. S. (2004). Is there opponent-orientation coding in the second-order channels of pattern vision? *Vision Res, 44*(27), 3145-3175. doi: 10.1016/j.visres.2004.07.018

Graham, N. V. (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): useful additions of the last 25 years. *Vision Res, 51*(13), 1397-1430. doi: 10.1016/j.visres.2011.02.007

Graziano, M. S., Andersen, R. A., & Snowden, R. J. (1994). Tuning of MST neurons to spiral motions. *J Neurosci, 14*(1), 54-67.

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron, 21*(1), 191-202.

Grzywacz, N. M., Watamaniuk, S. N., & McKee, S. P. (1995). Temporal coherence theory for the detection and measurement of visual motion. *Vision Res, 35*(22), 3183-3203.

Halgren, E., Mendola, J., Chong, C. D., & Dale, A. M. (2003). Cortical activation to illusory shapes as measured with magnetoencephalography. *Neuroimage, 18*(4), 1001-1009.

Hansen, B. C., & Hess, R. F. (2006). The role of spatial phase in texture segmentation and contour integration. *J Vis, 6*(5), 594-615. doi: 10.1167/6.5.5

Hayes, A. (2000). Apparent position governs contour-element binding by the visual system. *Proc Biol Sci, 267*(1450), 1341-1345. doi: 10.1098/rspb.2000.1148

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*(6598), 334-337. doi: 10.1038/383334a0

Heinrich, S. P. (2007). A primer on motion visual evoked potentials. *Doc Ophthalmol, 114*(2), 83-105. doi: 10.1007/s10633-006-9043-8

Heinrich, S. P., & Bach, M. (2003). Adaptation characteristics of steady-state motion visual evoked potentials. *Clin Neurophysiol, 114*(7), 1359-1366.

Heinrich, S. P., van der Smagt, M. J., Bach, M., & Hoffmann, M. B. (2004). Electrophysiological evidence for independent speed channels in human motion processing. *J Vis, 4*(6), 469-475. doi: 10:1167/4.6.6

Henry, G. H., Bishop, P. O., & Dreher, B. (1974). Orientation, axis and direction as stimulus parameters for striate cells. *Vision Res, 14*(9), 767-777.

Herrmann, C. S., & Bosch, V. (2001). Gestalt perception modulates early visual processing. *Neuroreport, 12*(5), 901-904.

Herrmann, C. S., & Knight, R. T. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neurosci Biobehav Rev, 25*(6), 465-476.

Hess, R., & Field, D. (1999). Integration of contours: new insights. *Trends Cogn Sci, 3*(12), 480-486.

Hess, R. F., & Dakin, S. C. (1997). Absence of contour linking in peripheral vision. *Nature, 390*(6660), 602-604. doi: 10.1038/37593

Hess, R. F., & Dakin, S. C. (1999). Contour integration in the peripheral field. *Vision Res, 39*(5), 947-959.

Hess, R. F., & Field, D. J. (1995). Contour integration across depth. *Vision Res, 35*(12), 1699-1711.

Hess, R. F., Hayes, A., & Field, D. J. (2003). Contour integration and cortical processing. *J Physiol Paris, 97*(2-3), 105-119. doi: 10.1016/j.jphysparis.2003.09.013

Hess, R. F., Hayes, A., & Kingdom, F. A. (1997). Integrating contours within and through depth. *Vision Res, 37*(6), 691-696.

Hess, R. F., & Ledgeway, T. (2003). The detection of direction-defined and speed-defined spatial contours: one mechanism or two? *Vision Res, 43*(5), 597-606.

Hesselmann, G., & Malach, R. (2011). The link between fMRI-BOLD activation and perceptual awareness is "stream-invariant" in the human visual system. *Cereb Cortex, 21*(12), 2829-2837. doi: 10.1093/cercor/bhr085

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proc Natl Acad Sci U S A, 95*(3), 781-787.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos Trans R Soc Lond B Biol Sci, 353*(1373), 1257-1270. doi: 10.1098/rstb.1998.0281

Hoffmann, M. B., Unsold, A. S., & Bach, M. (2001). Directional tuning of human motion adaptation as reflected by the motion VEP. *Vision Res, 41*(17), 2187-2194.

Holliday, I. E., & Meese, T. S. (2005). Neuromagnetic evoked responses to complex motions are greatest for expansion. *Int J Psychophysiol, 55*(2), 145-157. doi: 10.1016/j.ijpsycho.2004.07.009

Hopf, J. M., Boelmans, K., Schoenfeld, M. A., Luck, S. J., & Heinze, H. J. (2004). Attention to features precedes attention to locations in visual search: evidence from electromagnetic brain responses in humans. *J Neurosci, 24*(8), 1822-1832. doi: 10.1523/JNEUROSCI.3564-03.2004

Hopf, J. M., Vogel, E., Woodman, G., Heinze, H. J., & Luck, S. J. (2002). Localizing visual discrimination processes in time and space. *J Neurophysiol, 88*(4), 2088-2095.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol, 160*, 106-154.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J Physiol, 195*(1), 215-243.

Hubel, D. H., & Wiesel, T. N. (1977). Ferrier lecture. Functional architecture of macaque monkey visual cortex. *Proc R Soc Lond B Biol Sci, 198*(1130), 1-59.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci, 22*(16), 7195-7205. doi: 20026661

Jeffreys, D. A., & Axford, J. G. (1972). Source locations of pattern-specific components of human visual evoked potentials. II. Component of extrastriate cortical origin. *Exp Brain Res, 16*(1), 22-40.

Jewell, G., & McCourt, M. E. (2000). Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks. *Neuropsychologia, 38*(1), 93-110.

Johannes, S., Munte, T. F., Heinze, H. J., & Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Brain Res Cogn Brain Res, 2*(3), 189-205.

Johnston, A., & Scarfe, P. (2013). The role of the harmonic vector average in motion integration. *Front Comput Neurosci, 7*, 146. doi: 10.3389/fncom.2013.00146

Kaneoke, Y., Watanabe, S., & Kakigi, R. (2005). Human visual processing as revealed by magnetoencephalography. *Int Rev Neurobiol, 68*, 197-222. doi: 10.1016/S0074-7742(05)68008-7

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15*(4), 843-856.

Kastner, S., Nothdurft, H. C., & Pigarev, I. N. (1999). Neuronal responses to orientation and motion contrast in cat striate cortex. *Vis Neurosci, 16*(3), 587-600.

Kelly, S. P., Gomez-Ramirez, M., & Foxe, J. J. (2008). Spatial attention modulates initial afferent activity in human primary visual cortex. *Cereb Cortex, 18*(11), 2629-2636. doi: 10.1093/cercor/bhn022

Kennedy, J. M., & Domander, R. (1985). Shape and contour: the points of maximum change are least useful for recognition. *Perception, 14*(3), 367-370.

Khoe, W., Freeman, E., Woldorff, M. G., & Mangun, G. R. (2004). Electrophysiological correlates of lateral interactions in human visual cortex. *Vision Res, 44*(14), 1659-1673. doi: 10.1016/j.visres.2004.02.011

Kim, J., & Wilson, H. R. (1993). Dependence of plaid motion coherence on component grating directions. *Vision Res, 33*(17), 2479-2489.

Knebel, J. F., & Murray, M. M. (2012). Towards a resolution of conflicting models of illusory contour processing in humans. *Neuroimage, 59*(3), 2808-2817. doi: 10.1016/j.neuroimage.2011.09.031

Knierim, J. J., & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J Neurophysiol, 67*(4), 961-980.

Konen, C. S., & Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci, 11*(2), 224-231. doi: 10.1038/nn2036

Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proc Natl Acad Sci U S A, 90*(16), 7495-7497.

Kovacs, I., Papathomas, T. V., Yang, M., & Feher, A. (1996). When the brain changes its mind: interocular grouping during binocular rivalry. *Proc Natl Acad Sci U S A, 93*(26), 15508-15511.

Kremlacek, J., Kuba, M., Kubova, Z., & Chlubnova, J. (2004). Motion-onset VEPs to translating, radial, rotating and spiral stimuli. *Doc Ophthalmol, 109*(2), 169-175.

Kuba, M., & Kubova, Z. (1992). Visual evoked potentials specific for motion onset. *Doc Ophthalmol, 80*(1), 83-89.

Kubova, Z., & Kuba, M. (1995). Motion onset VEPs improve the diagnostics of multiple sclerosis and optic neuritis. *Sb Ved Pr Lek Fak Karlovy Univerzity Hradci Kralove Suppl, 38*(2), 89-93.

Kubova, Z., Kuba, M., Hubacek, J., & Vit, F. (1990). Properties of visual evoked potentials to onset of movement on a television screen. *Doc Ophthalmol, 75*(1), 67-72.

Kubova, Z., Kuba, M., Peregrin, J., & Novakova, V. (1996). Visual evoked potential evidence for magnocellular system deficit in dyslexia. *Physiol Res, 45*(1), 87-89.

Kubova, Z., Kuba, M., Spekreijse, H., & Blakemore, C. (1995). Contrast dependence of motion-onset and pattern-reversal evoked potentials. *Vision Res, 35*(2), 197-205.

Lagae, L., Maes, H., Raiguel, S., Xiao, D. K., & Orban, G. A. (1994). Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. *J Neurophysiol, 71*(5), 1597-1626.

Lagae, L., Raiguel, S., & Orban, G. A. (1993). Speed and direction selectivity of macaque middle temporal neurons. *J Neurophysiol, 69*(1), 19-39.

Lamme, V. A., & Spekreijse, H. (2000). Modulations of primary visual cortex activity representing attentive and conscious scene perception. *Front Biosci, 5*, D232-243.

Lamme, V. A., van Dijk, B. W., & Spekreijse, H. (1993). Contour from motion processing occurs in primary visual cortex. *Nature, 363*(6429), 541-543. doi: 10.1038/363541a0

Lamme, V. A., Van Dijk, B. W., & Spekreijse, H. (1994). Organization of contour from motion processing in primate visual cortex. *Vision Res, 34*(6), 721-735.

Lamme, V. A., Zipser, K., & Spekreijse, H. (2002). Masking interrupts figure-ground signals in V1. *J Cogn Neurosci, 14*(7), 1044-1053. doi: 10.1162/089892902320474490

Ledgeway, T., & Hess, R. F. (2002). Rules for combining the outputs of local motion detectors to define simple contours. *Vision Res, 42*(5), 653-659.

Ledgeway, T., & Hess, R. F. (2006). The spatial frequency and orientation selectivity of the mechanisms that extract motion-defined contours. *Vision Res, 46*(4), 568-578. doi: 10.1016/j.visres.2005.08.010

Ledgeway, T., Hess, R. F., & Geisler, W. S. (2005). Grouping local orientation and direction signals to extract spatial contours: empirical tests of "association field" models of contour integration. *Vision Res, 45*(19), 2511-2522. doi: 10.1016/j.visres.2005.04.002

Levi, D. M. (2008). Crowding--an essential bottleneck for object recognition: a mini-review. *Vision Res, 48*(5), 635-654. doi: 10.1016/j.visres.2007.12.009

Levi, D. M., & Waugh, S. J. (1996). Position acuity with opposite-contrast polarity features: evidence for a nonlinear collector mechanism for position acuity? *Vision Res, 36*(4), 573-588.

Levine, M. W., & McAnany, J. J. (2005). The relative capabilities of the upper and lower visual hemifields. *Vision Res, 45*(21), 2820-2830. doi: 10.1016/j.visres.2005.04.001

Li, W., & Gilbert, C. D. (2002). Global contour saliency and local colinear interactions. *J Neurophysiol, 88*(5), 2846-2856. doi: 10.1152/jn.00289.2002

Li, W., Piech, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron, 50*(6), 951-962. doi: 10.1016/j.neuron.2006.04.035

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Comput, 10*(4), 903-940.

Li, Z. (2000). Pre-attentive segmentation in the primary visual cortex. *Spat Vis, 13*(1), 25-50.

Lin, L. M., & Wilson, H. R. (1996). Fourier and non-Fourier pattern discrimination compared. *Vision Res, 36*(13), 1907-1918.

Loffler, G. (2008). Perception of contours and shapes: low and intermediate stage mechanisms. *Vision Res, 48*(20), 2106-2127. doi: 10.1016/j.visres.2008.03.006

Lopez-Calderon, J., & Luck, S. (2010). ERPLAB. Plug-in for EEGLAB.: In development at the Center for Mind and Brain, University of California at Davis.

Lorenceau, J., Giersch, A., & Series, P. (2005). Dynamic competition between contour integration and contour segmentation probed with moving stimuli. *Vision Res, 45*(1), 103-116. doi: 10.1016/j.visres.2004.07.033

Luck, S. J. (2005). *An introduction to the event-related potential technique (cognitive neuroscience).* Cambridge, MA: MIT Press.

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalogr Clin Neurophysiol, 75*(6), 528-542.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31*(3), 291-308.

Machilsen, B., Novitskiy, N., Vancleef, K., & Wagemans, J. (2011). Context modulates the ERP signature of contour integration. *PLoS One, 6*(9), e25151. doi: 10.1371/journal.pone.0025151

Maffei, L., & Campbell, F. W. (1970). Neurophysiological localization of the vertical and horizontal visual coordinates in man. *Science, 167*(3917), 386-387.

Malach, R., Amir, Y., Harel, M., & Grinvald, A. (1993). Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. *Proc Natl Acad Sci U S A, 90*(22), 10469-10473.

Malania, M., Herzog, M. H., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *J Vis, 7*(2), 1 1-7. doi: 10.1167/7.2.1

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32*(1), 4-18.

Mather, G., Pavan, A., Bellacosa Marotti, R., Campana, G., & Casco, C. (2013). Interactions between motion and form processing in the human visual system. *Front Comput Neurosci, 7*, 65. doi: 10.3389/fncom.2013.00065

Mather, G., Pavan, A., Bellacosa, R. M., & Casco, C. (2012). Psychophysical evidence for interactions between visual motion and form processing at the level of motion integrating receptive fields. *Neuropsychologia, 50*(1), 153-159. doi: 10.1016/j.neuropsychologia.2011.11.013

Mathes, B., Trenner, D., & Fahle, M. (2006). The electrophysiological correlate of contour integration is modulated by task demands. *Brain Res, 1114*(1), 98-112. doi: 10.1016/j.brainres.2006.07.068

Maunsell, J. H., Nealey, T. A., & DePriest, D. D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *J Neurosci, 10*(10), 3323-3334.

Maunsell, J. H., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annu Rev Neurosci, 10*, 363-401. doi: 10.1146/annurev.ne.10.030187.002051

Maurer, J. P., Heinrich, T. S., & Bach, M. (2004). Direction tuning of human motion detection determined from a population model. *Eur J Neurosci, 19*(12), 3359-3364. doi: 10.1111/j.0953-816X.2004.03419.x

May, K. A., & Hess, R. F. (2007a). Dynamics of snakes and ladders. *J Vis, 7*(12), 13 11-19. doi: 10.1167/7.12.13

May, K. A., & Hess, R. F. (2007b). Ladder contours are undetectable in the periphery: a crowding effect? *J Vis, 7*(13), 9 1-15. doi: 10.1167/7.13.9

May, K. A., & Hess, R. F. (2008). Effects of element separation and carrier wavelength on detection of snakes and ladders: implications for models of contour integration. *J Vis, 8*(13), 4 1-23. doi: 10.1167/8.13.4

McAnany, J. J., & Levine, M. W. (2007). Magnocellular and parvocellular visual pathway contributions to visual field anisotropies. *Vision Res, 47*(17), 2327-2336. doi: 10.1016/j.visres.2007.05.013

McGuire, B. A., Gilbert, C. D., Rivlin, P. K., & Wiesel, T. N. (1991). Targets of horizontal connections in macaque primary visual cortex. *J Comp Neurol, 305*(3), 370-392. doi: 10.1002/cne.903050303

McLean, J., & Palmer, L. A. (1989). Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in area 17 of cat. *Vision Res, 29*(6), 675-679.

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behav Brain Res, 6*(1), 57-77.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends In Neuroscience, 6*, 414-417.

Mitchison, G., & Crick, F. (1982). Long axons within the striate cortex: their distribution, orientation, and patterns of connection. *Proc Natl Acad Sci U S A, 79*(11), 3661-3665.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res Cogn Brain Res, 14*(1), 115-128.

Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nat Neurosci, 3*(12), 1322-1328. doi: 10.1038/81860

Motoyoshi, I., & Kingdom, F. A. (2007). Differential roles of contrast polarity reveal two streams of second-order visual processing. *Vision Res, 47*(15), 2047-2054. doi: 10.1016/j.visres.2007.03.015

Moulden, B. (1994). Collator units: second-stage orientational filters. *Higher-order processing in the visual system.* (Vol. 184, pp. 170-192): Ciba Foundation Symposium.

Movshon, J., Adelson, E., Gizzi, M., & Newsome, W. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass & C. Gross (Eds.), *Pattern Recognition Mechanisms* (pp. 117– 151). Rome: Vatican.

Muller, R., & Gopfert, E. (1988). The influence of grating contrast on the human cortical potential visually evoked by motion. *Acta Neurobiol Exp (Wars), 48*(5), 239-249.

Muller, R., Gopfert, E., Leineweber, M., & Greenlee, M. W. (2004). Effect of adaptation direction on the motion VEP and perceived speed of drifting gratings. *Vision Res, 44*(20), 2381-2392. doi: 10.1016/j.visres.2004.05.005

Nakamura, H., Kashii, S., Nagamine, T., Matsui, Y., Hashimoto, T., Honda, Y., & Shibasaki, H. (2003). Human V5 demonstrated by magnetoencephalography using random dot kinematograms of different coherence levels. *Neurosci Res, 46*(4), 423-433.

Nakayama, K., Silverman, G. H., MacLeod, D. I., & Mulligan, J. (1985). Sensitivity to shearing and compressive motion in random dots. *Perception, 14*(2), 225-238.

Nikolaev, A. R., & van Leeuwen, C. (2004). Flexibility in spatial and non-spatial feature grouping: an event-related potentials study. *Brain Res Cogn Brain Res, 22*(1), 13-25. doi: 10.1016/j.cogbrainres.2004.07.004

Norcia, A. M., Garcia, H., Humphry, R., Holmes, A., Hamer, R. D., & Orel-Bixler, D. (1991). Anomalous motion VEPs in infants and in infantile esotropia. *Invest Ophthalmol Vis Sci, 32*(2), 436-439.

Nugent, A. K., Keswani, R. N., Woods, R. L., & Peli, E. (2003). Contour integration in peripheral vision reduces gradually with eccentricity. *Vision Res, 43*(23), 2427-2437.

Nygard, G. E., Looy, T. V., & Wagemans, J. (2009). The influence of orientation jitter and motion on contour saliency and object identification. *Vision Res, 49*(20), 2475-2484. doi: 10.1016/j.visres.2009.08.002

Orban, G. A., Dupont, P., De Bruyn, B., Vogels, R., Vandenberghe, R., & Mortelmans, L. (1995). A motion area in human visual cortex. *Proc Natl Acad Sci U S A, 92*(4), 993-997.

Pack, C. C., Hunter, J. N., & Born, R. T. (2005). Contrast dependence of suppressive influences in cortical area MT of alert macaque. *J Neurophysiol, 93*(3), 1809-1815. doi: 10.1152/jn.00629.2004

Panis, S., De Winter, J., Vandekerckhove, J., & Wagemans, J. (2008). Identification of everyday objects on the basis of fragmented outline versions. *Perception, 37*(2), 271-289.

Panis, S., & Wagemans, J. (2009). Time-course contingencies in perceptual organization and identification of fragmented object outlines. *J Exp Psychol Hum Percept Perform, 35*(3), 661-687. doi: 10.1037/a0013547

Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *J Neurophysiol, 82*(5), 2490-2502.

Pavan, A., Casco, C., Mather, G., Bellacosa, R. M., Cuturi, L. F., & Campana, G. (2011). The effect of spatial orientation on detecting motion trajectories in noise. *Vision Res, 51*(18), 2077-2084. doi: 10.1016/j.visres.2011.08.001

Pavan, A., Marotti, R. B., & Mather, G. (2013). Motion-form interactions beyond the motion integration level: evidence for interactions between orientation and optic flow signals. *J Vis, 13*(6), 16. doi: 10.1167/13.6.16

Pazo-Alvarez, P., Amenedo, E., & Cadaveira, F. (2004). Automatic detection of motion direction changes in the human brain. *Eur J Neurosci, 19*(7), 1978-1986. doi: 10.1111/j.1460-9568.2004.03273.x

Pei, F., Pettet, M. W., Vildavski, V. Y., & Norcia, A. M. (2005). Event-related potentials show configural specificity of global form processing. *Neuroreport, 16*(13), 1427-1430.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis, 10*(4), 437-442.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J Vis, 4*(12), 1136-1169. doi: 10:1167/4.12.12

Phillips, F., & Todd, J. T. (2010). Texture discrimination based on global feature alignments. *J Vis, 10*(6), 6. doi: 10.1167/10.6.6

Pitts, M. A., Martinez, A., Brewer, J. B., & Hillyard, S. A. (2011). Early stages of figure-ground segregation during perception of the face-vase. *J Cogn Neurosci, 23*(4), 880-895. doi: 10.1162/jocn.2010.21438

Pitzalis, S., Sdoia, S., Bultrini, A., Committeri, G., Di Russo, F., Fattori, P., . . . Galati, G. (2013). Selectivity to translational egomotion in human brain motion areas. *PLoS One, 8*(4), e60241. doi: 10.1371/journal.pone.0060241

Polat, U. (1999). Functional architecture of long-range perceptual interactions. *Spat Vis, 12*(2), 143-162.

Polat, U., & Bonneh, Y. (2000). Collinear interactions and contour integration. *Spat Vis, 13*(4), 393-401.

Polat, U., & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Res, 34*(1), 73-78.

Prins, N., Kingdom, F. A., & Hayes, A. (2007). Detecting low shape-frequencies in smooth and jagged contours. *Vision Res, 47*(18), 2390-2402. doi: 10.1016/j.visres.2007.06.006

Probst, T., Plendl, H., Paulus, W., Wist, E. R., & Scherg, M. (1993). Identification of the visual motion area (area V5) in the human brain by dipole source analysis. *Exp Brain Res, 93*(2), 345-351.

Rademacher, J., Caviness, V. S., Jr., Steinmetz, H., & Galaburda, A. M. (1993). Topographical variation of the human primary cortices: implications for neuroimaging, brain mapping, and neurobiology. *Cereb Cortex, 3*(4), 313-329.

Raiguel, S., Van Hulle, M. M., Xiao, D. K., Marcar, V. L., & Orban, G. A. (1995). Shape and spatial distribution of receptive fields and antagonistic motion surrounds in the middle temporal area (V5) of the macaque. *Eur J Neurosci, 7*(10), 2064-2082.

Rainville, S. J., & Wilson, H. R. (2005). Global shape coding for motion-defined radial-frequency contours. *Vision Res, 45*(25-26), 3189-3201. doi: 10.1016/j.visres.2005.06.033

Rauss, K. S., Pourtois, G., Vuilleumier, P., & Schwartz, S. (2009). Attentional load modifies early activity in human primary visual cortex. *Hum Brain Mapp, 30*(5), 1723-1733. doi: 10.1002/hbm.20636

Reid, R. C., Soodak, R. E., & Shapley, R. M. (1987). Linear mechanisms of directional selectivity in simple cells of cat striate cortex. *Proc Natl Acad Sci U S A, 84*(23), 8740-8744.

Ritter, W., Simson, R., Vaughan, H. G., Jr., & Friedman, D. (1979). A brain event related to the making of a sensory discrimination. *Science, 203*(4387), 1358-1361.

Robol, V., Casco, C., & Dakin, S. C. (2012). The role of crowding in contextual influences on contour integration. *J Vis, 12*(7), 3. doi: 10.1167/12.7.3

Robol, V., Grassi, M., & Casco, C. (2013). Contextual influences in texture-segmentation: distinct effects from elements along the edge and in the texture-region. *Vision Res, 88*, 1-8. doi: 10.1016/j.visres.2013.05.010

Rockland, K. S., & Lund, J. S. (1982). Widespread periodic intrinsic connections in the tree shrew visual cortex. *Science, 215*(4539), 1532-1534.

Rockland, K. S., & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *J Comp Neurol, 216*(3), 303-318. doi: 10.1002/cne.902160307

Rockland, K. S., Lund, J. S., & Humphrey, A. L. (1982). Anatomical binding of intrinsic connections in striate cortex of tree shrews (Tupaia glis). *J Comp Neurol, 209*(1), 41-58. doi: 10.1002/cne.902090105

Roelfsema, P. R., Lamme, V. A., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature, 395*(6700), 376-381. doi: 10.1038/26475

Roncato, S., & Casco, C. (2006). Illusory boundary interpolation from local association field. *Spat Vis, 19*(6), 581-603.

Roncato, S., & Casco, C. (2009). A new "tilt" illusion reveals the relation between border ownership and border binding. *J Vis, 9*(6), 14 11-10. doi: 10.1167/9.6.14

Rossi, V., & Pourtois, G. (2012). State-dependent attention modulation of human primary visual cortex: a high density ERP study. *Neuroimage, 60*(4), 2365-2378. doi: 10.1016/j.neuroimage.2012.02.007

Rossion, B., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guerit, J. M. (1999). Spatio-temporal localization of the face inversion effect: an event-related potentials study. *Biol Psychol, 50*(3), 173-189.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behav Cogn Neurosci Rev, 1*(1), 63-75.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychol Sci, 13*(3), 250-257.

Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004). How parallel is visual processing in the ventral pathway? *Trends Cogn Sci, 8*(8), 363-370. doi: 10.1016/j.tics.2004.06.003

Samonds, J. M., Zhou, Z., Bernard, M. R., & Bonds, A. B. (2006). Synchronous activity in cat visual cortex encodes collinear and cocircular contours. *J Neurophysiol, 95*(4), 2602-2616. doi: 10.1152/jn.01070.2005

Sarti, A., Citti, G., & Petitot, J. (2009). Functional geometry of the horizontal connectivity in the primary visual cortex. *J Physiol Paris, 103*(1-2), 37-45. doi: 10.1016/j.jphysparis.2009.05.004

Schellart, N. A., Trindade, M. J., Reits, D., Verbunt, J. P., & Spekreijse, H. (2004). Temporal and spatial congruence of components of motion-onset evoked responses investigated by whole-head magneto-electroencephalography. *Vision Res, 44*(2), 119-134.

Schendan, H. E., Ganis, G., & Kutas, M. (1998). Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology, 35*(3), 240-251.

Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976). Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *J Neurophysiol, 39*(6), 1288-1319.

Schmidt, K. E., Goebel, R., Lowel, S., & Singer, W. (1997). The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex. *Eur J Neurosci, 9*(5), 1083-1089.

Schoenfeld, M. A., Woldorff, M., Duzel, E., Scheich, H., Heinze, H. J., & Mangun, G. R. (2003). Form-from-motion: MEG evidence for time course and processing sequence. *J Cogn Neurosci, 15*(2), 157-172. doi: 10.1162/089892903321208105

Scholte, H. S., Jolij, J., Fahrenfort, J. J., & Lamme, V. A. (2008). Feedforward and recurrent processing in scene segmentation: electroencephalography and functional magnetic resonance imaging. *J Cogn Neurosci, 20*(11), 2097-2109. doi: 10.1162/jocn.2008.20142

Schumacher, J. F., Quinn, C. F., & Olman, C. A. (2011). An exploration of the spatial scale over which orientation-dependent surround effects affect contour detection. *J Vis, 11*(8), 12. doi: 10.1167/11.8.12

Senkowski, D., Rottger, S., Grimm, S., Foxe, J. J., & Herrmann, C. S. (2005). Kanizsa subjective figures capture visual spatial attention: evidence from electrophysiological and behavioral data. *Neuropsychologia, 43*(6), 872-886. doi: 10.1016/j.neuropsychologia.2004.09.010

Shpaner, M., Molholm, S., Forde, E., & Foxe, J. J. (2013). Disambiguating the roles of area V1 and the lateral occipital complex (LOC) in contour integration. *Neuroimage, 69*, 146-156. doi: 10.1016/j.neuroimage.2012.11.023

Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: natural scenes and Gestalt rules. *Proc Natl Acad Sci U S A, 98*(4), 1935-1940. doi: 10.1073/pnas.031571498

Skrandies, W. (1987). *The upper and lower visual field of man: Electrophysiological and functional differences.* Berlin Heidelberg: Springer.

Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M., & Hennig, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *J Neurosci, 18*(10), 3816-3830.

Snowden, R. J., & Braddick, O. J. (1989). The combination of motion signals over time. *Vision Res, 29*(11), 1621-1630.

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *J Neurosci, 11*(9), 2768-2785.

Spillmann, L., & Werner, J. S. (1996). Long-range interactions in visual perception. *Trends Neurosci, 19*(10), 428-434.

Stettler, D. D., Das, A., Bennett, J., & Gilbert, C. D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron, 36*(4), 739-750.

Straube, S., & Fahle, M. (2010). The electrophysiological correlate of saliency: evidence from a figure-detection task. *Brain Res, 1307*, 89-102. doi: 10.1016/j.brainres.2009.10.043

Straube, S., Grimsen, C., & Fahle, M. (2010). Electrophysiological correlates of figure-ground segregation directly reflect perceptual saliency. *Vision Res, 50*(5), 509-521. doi: 10.1016/j.visres.2009.12.013

Tanaka, K., Fukada, Y., & Saito, H. A. (1989). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol, 62*(3), 642-656.

Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol, 62*(3), 626-641.

Tanskanen, T., Saarinen, J., Parkkonen, L., & Hari, R. (2008). From local to global: Cortical dynamics of contour integration. *J Vis, 8*(7), 15 11-12. doi: 10.1167/8.7.15

Thiele, A., Dobkins, K. R., & Albright, T. D. (2001). Neural correlates of chromatic motion perception. *Neuron, 32*(2), 351-358.

Thomas, N. A., & Elias, L. J. (2011). Upper and lower visual field differences in perceptual asymmetries. *Brain Res, 1387*, 108-115. doi: 10.1016/j.brainres.2011.02.063

Todd, J. T., & Norman, J. F. (1995). The effects of spatiotemporal integration on maximum displacement thresholds for the detection of coherent motion. *Vision Res, 35*(16), 2287-2302.

Tootell, R. B., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Liu, A. K., Reppas, J. B., . . . Dale, A. M. (1997). Functional analysis of V3A and related areas in human visual cortex. *J Neurosci, 17*(18), 7060-7078.

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., . . . Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci, 15*(4), 3215-3230.

Torriente, I., Valdes-Sosa, M., Ramirez, D., & Bobes, M. A. (1999). Visual evoked potentials related to motion-onset are modulated by attention. *Vision Res, 39*(24), 4122-4139.

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

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

Ungerleider, L. G., & Pasternak, T. (2004). Ventral and dorsal cortical processing streams. In W. J. S. Chalupa L. M. (Ed.), *The Visual Neurosciences.* Cambridge, MA: MIT Press.

Ursino, M., & La Cara, G. E. (2004). Comparison of different models of orientation selectivity based on distinct intracortical inhibition rules. *Vision Res, 44*(14), 1641-1658. doi: 10.1016/j.visres.2004.02.005

Usher, M., Bonneh, Y., Sagi, D., & Herrmann, M. (1999). Mechanisms for spatial integration in visual detection: a model based on lateral interactions. *Spat Vis, 12*(2), 187-209.

Uttal, W. R., Spillmann, L., Sturzel, F., & Sekuler, A. B. (2000). Motion and shape in common fate. *Vision Res, 40*(3), 301-310.

Van Essen, D. C., & Maunsell, J. H. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends In Neuroscience, 6*, 370-375.

Van Oostende, S., Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. A. (1997). The kinetic occipital (KO) region in man: an fMRI study. *Cereb Cortex, 7*(7), 690-701.

Vancleef, K., Wagemans, J., & Humphreys, G. W. (2013). Impaired texture segregation but spared contour integration following damage to right posterior parietal cortex. *Exp Brain Res, 230*(1), 41-57. doi: 10.1007/s00221-013-3629-7

Verghese, P., McKee, S. P., & Grzywacz, N. M. (2000). Stimulus configuration determines the detectability of motion signals in noise. *J Opt Soc Am A Opt Image Sci Vis, 17*(9), 1525-1534.

Verstraten, F. A., Fredericksen, R. E., & van de Grind, W. A. (1994). Movement aftereffect of bi-vectorial transparent motion. *Vision Res, 34*(3), 349-358.

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology, 37*(2), 190-203.

Vreven, D., & Verghese, P. (2002). Integration of speed signals in the direction of motion. *Percept Psychophys, 64*(6), 996-1007.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). Perceptual grouping and figure–ground organization. In J. Wagemans (Ed.), *A century of Gestalt psychology in visual perception: I.*

Wallach, H. (1935). Uber visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung, 20*, 325-380.

Watt, R., Ledgeway, T., & Dakin, S. C. (2008). Families of models for gabor paths demonstrate the importance of spatial adjacency. *J Vis, 8*(7), 23 21-19. doi: 10.1167/8.7.23

Watt, R. J., & Andrews, D. P. (1982). Contour curvature analysis: hyperacuities in the discrimination of detailed shape. *Vision Res, 22*(4), 449-460.

Wattam-Bell, J. (1994). Coherence thresholds for discrimination of motion direction in infants. *Vision Res, 34*(7), 877-883.

Wehrhahn, C., & Dresp, B. (1998). Detection facilitation by collinear stimuli in humans: dependence on strength and sign of contrast. *Vision Res, 38*(3), 423-428.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nat Neurosci, 5*(6), 598-604. doi: 10.1038/nn858

Weisstein, N., Maguire, W., & Berbaum, K. (1977). A phantom-motion aftereffect. *Science, 198*(4320), 955-958.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt, II. Psychologische Forschung. (W. D. Ellis, Trans.) *A source book of Gestalt psychology.* London: Routledge & Kegan Paul Ltd.

Wertheimer, M. (1938). Über Gestalttheorie (W. D. Ellis, Trans.) *A source book of Gestalt psychology.* (pp. 1-11). London: U. K.: Routledge & Kegan Paul Ltd.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends Cogn Sci, 6*(5), 211-216.

Williams, C. B., & Hess, R. F. (1998). Relationship between facilitation at threshold and suprathreshold contour integration. *J Opt Soc Am A Opt Image Sci Vis, 15*(8), 2046-2051.

Williams, D., Phillips, G., & Sekuler, R. (1986). Hysteresis in the perception of motion direction as evidence for neural cooperativity. *Nature, 324*(6094), 253-255. doi: 10.1038/324253a0

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis Neurosci, 9*(1), 79-97.

Wilson, H. R., & Kim, J. (1994). Perceived motion in the vector sum direction. *Vision Res, 34*(14), 1835-1842.

Wist, E. R., Gross, J. D., & Niedeggen, M. (1994). Motion aftereffects with random-dot chequerboard kinematograms: relation between psychophysical and VEP measures. *Perception, 23*(10), 1155-1162.

Wokke, M. E., Scholte, H. S., & Lamme, V. A. (2014). Opposing Dorsal/Ventral Stream Dynamics during Figure-ground Segregation. *J Cogn Neurosci, 26*(2), 365-379. doi: 10.1162/jocn_a_00497

Wokke, M. E., Sligte, I. G., Steven Scholte, H., & Lamme, V. A. (2012). Two critical periods in early visual cortex during figure-ground segregation. *Brain Behav, 2*(6), 763-777. doi: 10.1002/brb3.91

Yang, Y., & Blake, R. (1994). Broad tuning for spatial frequency of neural mechanisms underlying visual perception of coherent motion. *Nature, 371*(6500), 793-796. doi: 10.1038/371793a0

Yen, S. C., & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision Res, 38*(5), 719-741.

Yu, C., & Levi, D. M. (1997). Spatial facilitation predicted with end-stopped spatial filters. *Vision Res, 37*(22), 3117-3127.

Yuille, A. L., & Grzywacz, N. M. (1988). A computational theory for the perception of coherent visual motion. *Nature, 333*(6168), 71-74. doi: 10.1038/333071a0

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *J Neurosci, 11*(3), 641-649.

Zhaoping, L., & May, K. A. (2007). Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex. *PLoS Comput Biol, 3*(4), e62. doi: 10.1371/journal.pcbi.0030062