

# In Search of Texture Integration in the Early Visual Cortex

Sander Foekema

1280945

April 2010

Master Thesis  
Artificial Intelligence  
Department of Artificial Intelligence  
University of Groningen, the Netherlands

External Supervisor:

Dr. F.W. Cornelissen (Laboratory for Experimental Ophthalmology,  
University Medical Center Groningen, University of Groningen, the  
Netherlands)

External Supervisor 2:

Drs. K.V. Haak (Laboratory for Experimental Ophthalmology, Uni-  
versity Medical Center Groningen, University of Groningen, the  
Netherlands)

Internal Supervisor:

Drs. Gert Kootstra (Artificial Intelligence, University of Groningen)



university of  
 groningen

faculty of mathematics and  
 natural sciences

artificial intelligence



## **Voorwoord**

Voor u ligt het resultaat van mijn afstudeeronderzoek voor de master Artificial Intelligence, Rijksuniversiteit Groningen. Het onderzoek richt zich op onderliggende mechanismen van textuurverwerking in de visuele cortex. Hiervoor is gebruik gemaakt van een geavanceerde techniek om indirect hersenactiviteit te meten, namelijk fMRI, en is mogelijk gemaakt door de afdeling Laboratory for Experimental Ophthalmology en BCN NeuroImaging Center, Universitair Medisch Centrum Groningen (UMCG). Een scriptie is niet alleen een individueel proces en daarom wil ik graag wat mensen bedanken.

Allereerst, Frans Cornelissen bedankt voor het bewerkstelligen van de kans om een fMRI studie te kunnen doen en de begeleiding die daarbij komt kijken. Verder stond ook Koen Haak aan de basis van het vormen van mijn afstudeeronderzoek. Hiernaast is Koen dagelijks betrokken geweest bij het onderzoek. Dit heb ik op persoonlijk en professioneel vlak als zeer prettig ervaren en heb dan ook een hoop geleerd van zijn visie op het onderzoek en wetenschappelijk onderzoek in het algemeen. Bedankt! Verder wil ik graag mijn interne begeleider, Gert Kootstra, bedanken die, ondanks weinig raakvlakken met het thema van het onderzoek, veel interesse heeft getoond en mijn scriptie tot een goed einde heeft gebracht.

Een scriptie komt niet alleen tot stand met hulp van mensen die nauw betrokken zijn bij een onderzoek. Ook vrienden en collega's zijn belangrijk voor onder andere de gezelligheid, afleiding en werksfeer. Daarom wou ik mijn dank uitten aan mijn collega's, Remco, Jan-Bernard, Richard, Aditya, Ronald en Erik voor de goede werksfeer op de kamer op het NiC en voor de professionele input in het onderzoek. Vooral de laatste maanden van het schrijven van mijn scriptie waren zwaar en heb dan ook veel steun gehad van vrienden. In het bijzonder, Dennis voor de afspraak om elke dag een sigaretje en een bakje thee te drinken in de vroege morgen. Verder wil ik uiteraard mijn ouders en vriendin bedanken voor het medeleven en steun die ze mij hebben gegeven.

Wanneer ik terugkijk op het onderzoek en het schrijven van de scriptie zou ik nu veel dingen anders doen. Maar dat is nou juist iets waar je veel van leert. Wellicht kan er door de leuke resultaten van mijn scriptie nog een artikel worden geschreven. Alleen door de omvang en enkele tegenslagen is hier (nog) geen tijd voor geweest. Desondanks voelt de tijd die ik heb besteed aan het onderzoek als een verrijking van mijn persoonlijke en professionele ontwikkeling. Nu is er een nieuwe fase in mijn leven aangebroken en daar hoort natuurlijk een leuke baan bij. Ben benieuwd wat de toekomst mij zal brengen...

**Abstract**

In normal vision, objects, faces, and letters are centered onto the retina. One reason for this is that compared to foveal vision, peripheral vision lags the acuity to extract details. Another reason is a phenomenon called crowding. The crowding effect entails excessive grouping of visual features and is bound to a critical region that is linearly proportional to the viewing eccentricity. After decades of research, inappropriate 'hardwired' integration of afferent visual information in the visual cortex is the prevailing explanation of crowding. The aim of this research is finding the neural correlates of 'integration fields' in the early visual cortex using functional Magnetic Resonance Imaging (fMRI). The study comprised two fMRI experiments. Firstly, a standard visual field mapping procedure was performed that allowed the delineation of visual cortical regions. Secondly, fMRI responses were measured to the presentation of orientation-based texture stimuli where second-order orientational differences induce the percept of an illusory contour. This enabled a signature plot of neuronal integration processes in the early visual cortex as a function of cortical distance to the illusory contour. The resulting signature revealed integration processes as early as V1 and are compared with a biologically plausible model of crowding.

Keywords: texture integration, early visual cortex, crowding, fMRI

# Contents

<b>Chapter 1: Introduction</b>	<b>1</b>
1.1 Integration Fields . . . . .	3
1.2 Signature Integration Fields . . . . .	5
1.3 Thesis Outline . . . . .	7
<b>Chapter 2: Human Vision &amp; Crowding</b>	<b>9</b>
2.1 Pre-Cortical . . . . .	9
2.2 Visual Cortex . . . . .	11
2.3 Crowding . . . . .	13
2.3.1 Underlying Mechanisms of Crowding . . . . .	15
<b>Chapter 3: Visual Field Mapping</b>	<b>17</b>
3.1 Visual Field Mapping Technique . . . . .	17
3.1.1 Traveling-wave method . . . . .	18
3.1.2 Population receptive field method . . . . .	20
3.2 Materials and Methods . . . . .	21
3.2.1 Subjects . . . . .	21
3.2.2 Magnetic Resonance Imaging . . . . .	21
3.2.3 Anatomical preprocessing . . . . .	23
3.2.4 Functional preprocessing . . . . .	24
3.2.5 Gray and white matter segmentation . . . . .	24
3.2.6 Visualizing the visual cortex . . . . .	25
3.2.7 Stimulus presentation . . . . .	25
3.2.8 Stimulus description . . . . .	26
3.3 Results . . . . .	28

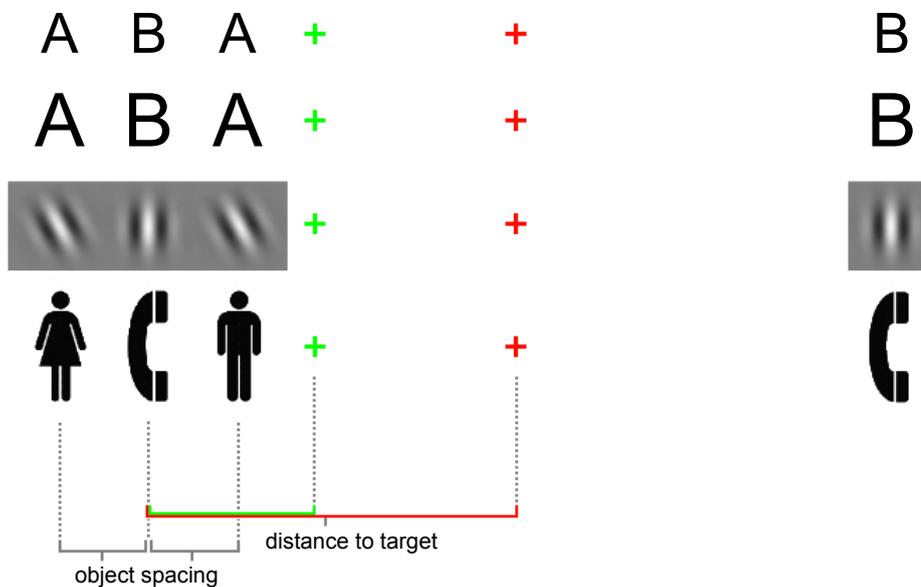
<b>Chapter 4: Signature Integration Fields: an fMRI approach</b>	<b>34</b>
4.1 Signature of Neural Processes . . . . .	35
4.1.1 Signature integration fields . . . . .	35
4.2 Materials and Methods . . . . .	36
4.2.1 Acquisition and preprocessing . . . . .	36
4.2.2 Stimulus description . . . . .	37
4.2.3 Center-Surround Localizer . . . . .	45
4.2.4 Signature mapping . . . . .	45
4.3 Results . . . . .	46
<b>Chapter 5: Signature Integration Fields: a Theoretical Approach</b>	<b>53</b>
5.1 The Crowding Model . . . . .	53
5.1.1 Stimulus encoding . . . . .	54
5.1.2 Stimulus integration . . . . .	57
5.2 Materials and Methods . . . . .	58
5.2.1 Stimulus description . . . . .	58
5.2.2 Signature mapping . . . . .	59
5.3 Results . . . . .	60
<b>Chapter 6: Discussion</b>	<b>62</b>
6.1 Empirical Signature of Integration fields . . . . .	63
6.2 Theoretical Signature of Integration Fields . . . . .	65
6.3 Comparison Empirical and Theoretical Signature . . . . .	66
6.4 General discussion . . . . .	67
<b>References</b>	<b>69</b>
<b>Appendix A: Brain Imaging</b>	<b>77</b>
A.1 The (f)MRI technique . . . . .	77
A.1.1 Anatomical Imaging . . . . .	79
A.1.2 Functional Imaging . . . . .	79
A.2 (f)MRI Procedure . . . . .	81
<b>Appendix B: Anatomical Terminology</b>	<b>83</b>
<b>Appendix C: Supplementary</b>	<b>85</b>
C.1 Projected Amplitude . . . . .	85
C.2 Individual Data . . . . .	85

## Introduction

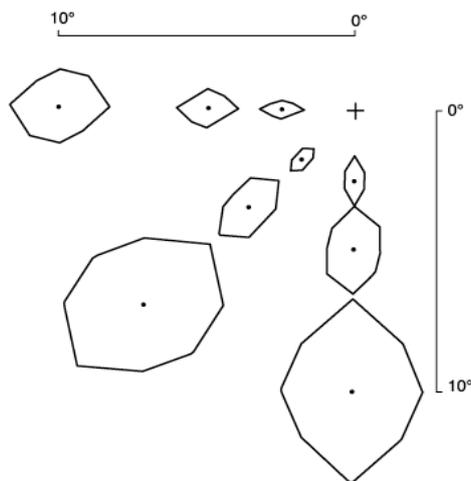
In normal vision, objects, faces, and letters are centered onto the retina. One reason for this is that compared to foveal vision, peripheral vision lags the acuity to extract details. Another reason is a phenomenon called *crowding* (Levi, 2008). Whereas visual acuity is reflected in our brain by cortical magnification, i.e. from fovea to periphery the number of cells involved in visual processing decreases exponentially, the neural correlates of crowding are still unknown. However, it is hypothesized that ‘hardwired’ integration of visual features over an inappropriate large area in peripheral vision results in the crowding effect. The aim of this research is finding the neural correlates of ‘hardwired’ visual feature integration in the early visual cortex using functional Magnetic Resonance Imaging (fMRI). Furthermore, a virtual fMRI experiment is performed with a biological plausible model of crowding to predict neural activity in the early visual cortex.

First observed by Korte (1923), the crowding phenomenon occurs when letters are located too close to each other to be identified, imposing a major impairment in letter recognition. For instance, when directing the center of gaze on the upper red fixation cross in figure 1.1, the target letter *B* on the right can easily be identified. In contrast, recognition of the target letter *B* on the left side of the figure is impaired. Subjects report seeing a jumbled precept of the crowded letters. Crowding is not only present in letter tasks, but also in, for instance, orientation, size, saturation, and hue identification tasks (van den Berg et al., 2007) (see other examples in figure 1.1). Remarkably, the center-to-center distance between objects need to exceed a critical spacing of roughly half of the visual angle distance to the target (eccentricity) to allow identification (Bouma, 1970). This is referred to as Bouma’s law and is now believed to be the defining property of crowding because the spatial extent of critical spacing is very robust and size- and type invariant.

New insights reveal that critical spacing can be extended to a two-dimensional



**Figure 1.1:** Examples of the crowding effect. When directing the center of gaze onto the red fixation cross, with a viewing distance of 15 cm, the eccentricity between the red fixation cross and the target objects on the left is 6 degrees. Then, the object spacing between target and flankers is less than the critical spacing and therefore the crowding effect prevents recognition of the target (the task for the third row is angle identification). In contrast, when fixating on the green fixation cross, the distance between the target and flankers exceed the critical spacing and therefore the target is not crowded. The objects on the right is not surrounded by distractors and are therefore not crowded.



**Figure 1.2:** The critical region of crowding. Here, the outerbounds of critical spacing in every direction is shown. The critical spacing is proportional to approximately  $0,5 \times$  eccentricity in radial and  $0,1 \times$  eccentricity in tangential direction. Objects that fall within the critical region are crowded. (source: Pelli (2008)).

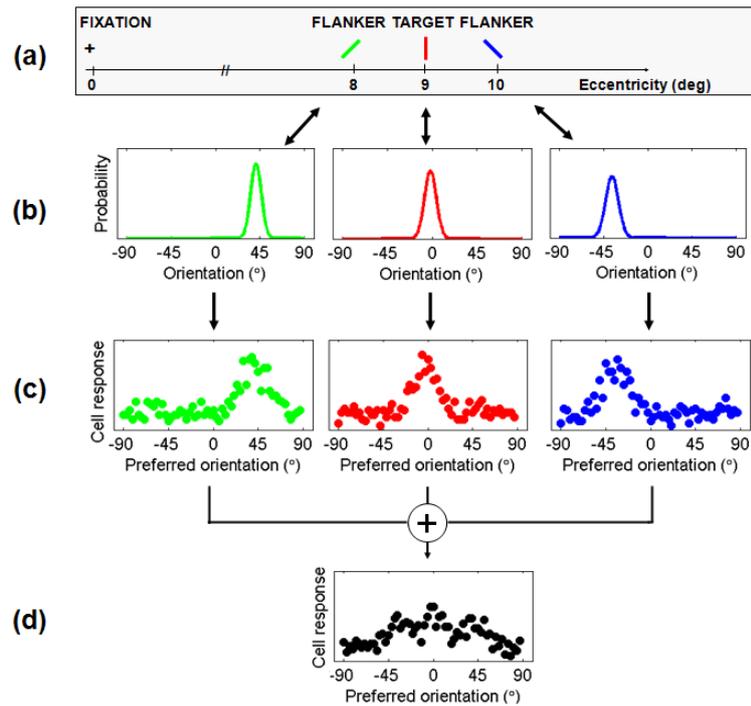
visual critical region wherein crowding takes place (Toet & Levi, 1992; Pelli, 2008), see figure 1.2 . Despite the inability to accurately recognize a target surrounded by distractors, Parkes et al. (2001) have shown, in an orientation discrimination task, that subjects readily report the average orientation of the target and flankers that fall within the critical region. Thus, it seems that visual features that fall within the critical region are integrated over an inappropriate large area in especially peripheral vision, imposing a major impairment in object recognition. However, texture segmentation and segregation seems to rely on processes that unravel statistical properties of the visual scene. In other words, the integration of visual features, explaining the crowding effect, is a process that unravels statistical properties and therefore plays an important role in texture segmentation and segregation.

## 1.1 Integration Fields

After decades of research, ‘hardwired’ integration of visual features over an inappropriate large area, i.e. critical region, appears to be the prevailing explanation of crowding (Pelli et al., 2004). That is, neuronal populations in the visual cortex pool similar afferent visual features that fall within area that is similar to the critical region of the crowding effect. Therefore, the spatial extent in the visual field wherein neurons pool visual features are addressed as ‘integration fields’ throughout this thesis.

In van den Berg et al. (2010) a biologic plausible quantitative model is proposed for the neuronal processes involved in crowding (see figure 1.3 for a graphical illustration of the model). The model consists of two important steps. First, neuronal signals in response to single variables, bars or edges, are computed based on population coding principles (Pouget et al., 2000; Zemel et al., 1998). Second, feature integration is modeled as the pooling of neuronal responses of the single variables within an integration field, where integration fields are modeled as a two-dimensional Gaussian weighted overlay function. Thus, the pooling of orientational features is reflected as compulsory averaging inside an integration field and results in probability distributions, i.e. ‘statistical summaries’, of angle information.

The ‘statistical summary’ produced by integration fields are important in visual processes involving texture analysis and segregation. For example, in figure 1.4a, a texture-like image is constructed of many oriented Gabor patches. Here, the orientational properties of the Gabor patches in the disk and annulus are drawn from distinct normal distributions, where  $\mu(\theta)_{disk} = -22.5^\circ$ ,  $\mu(\theta)_{annulus} = 22.5^\circ$



**Figure 1.3:** *Biological Plausible Model of Crowding.* The model consists of several steps. (a) The stimulus consist of three orientational bars. (b) The encoding of angle information of the individual stimuli are subjected to neural noise which causes uncertainty. (c) Then, the first layer of the model encodes the probability distribution of the individual stimuli into neural spike rates according to population coding techniques. (d) Then, an integration field that is centered onto the middle bar, pools the encoded signals. This result in a probability distribution of the expected target. The probability distribution has no clear distinction of the targets orientation, i.e. the target is crowded with the surrounding flankers. This results in a jumbled representation of the oriented bars at a certain location. (source: van den Berg et al. (2010))

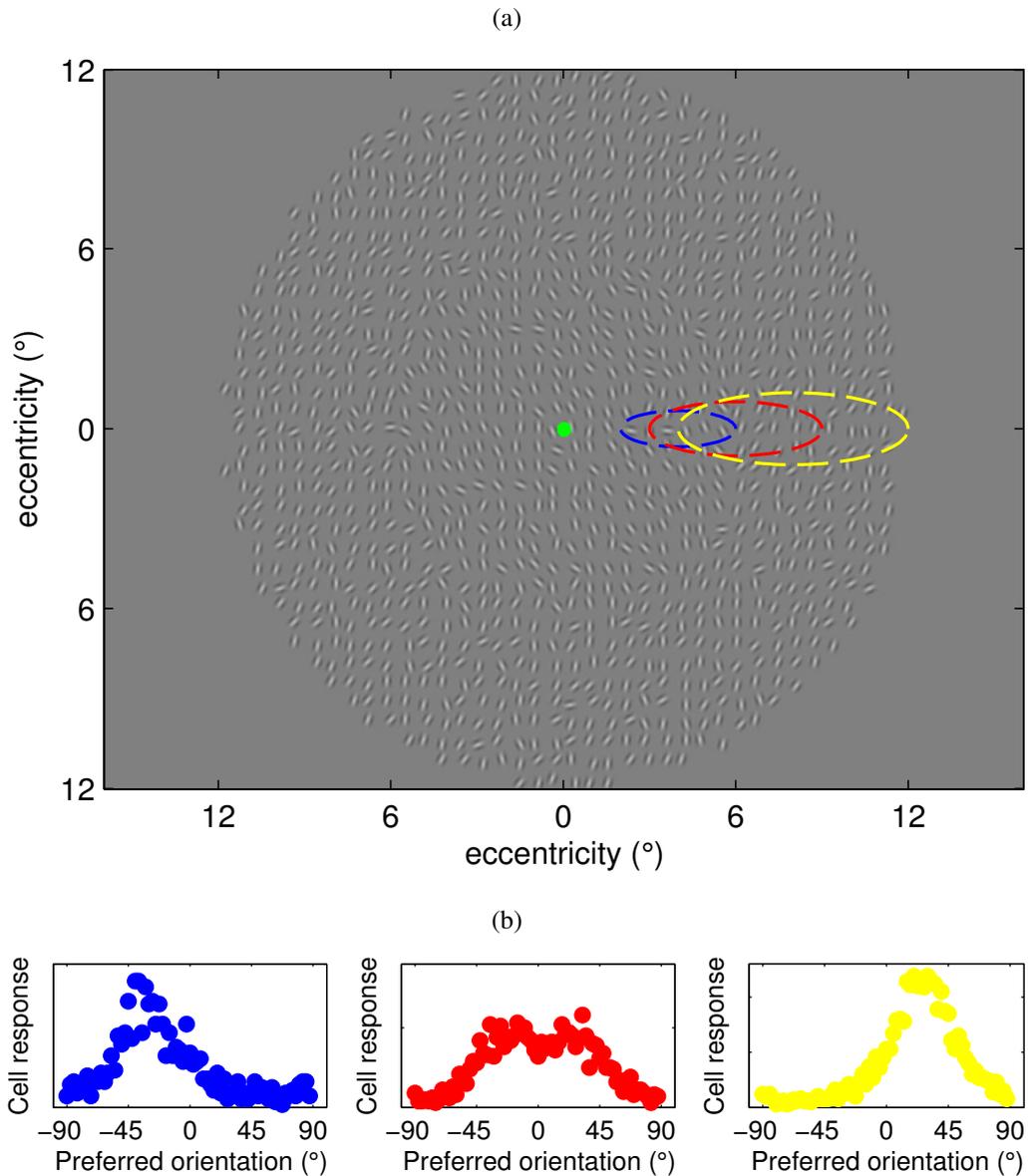
and  $VAR(\theta)_{disk} = VAR(\theta)_{annulus} = 20^\circ$ . An illusory contour is perceived at the transition between the disk and annulus because there is a substantially second order statistical difference between the orientational properties of the disk and annulus. Second order-statistics were used because these are more discriminating in human texture analysis than mere first-order statistics and it seems to be unessential to use statistics of higher order than of order 2 (Julesz, 1975). According to the crowding model, integration fields near the illusory contour produce a probability distribution that reflects the orientational properties of both the disk and annulus. In contrast, the net responses of integration fields encapsulating only the disk or annulus results in a more profound probability distribution (see probability distributions in figure 1.4b).

The hypothesis in this thesis is that neural correlates of integration processes in the early visual cortex (V1 and V2) can be found by measuring brain activity with functional Magnetic Resonance Imaging (fMRI) in combination with texture-like stimuli that induce the percept of an illusory contour. In other words, differences in the net-responses of integration fields can be measured in the early visual cortex using fMRI. Furthermore, we hypothesize that the biological plausible crowding model can be used to predict brain activity of the experimental paradigm in the fMRI experiment. A comparison between the fMRI analysis and model predictions should provide an understanding and validity of the neural correlates of integration fields.

## 1.2 Signature Integration Fields

The study comprises two fMRI experiments and one theoretical study. Firstly, a standard visual field mapping procedure will be performed (high contrast rotating wedges, contracting/expanding rings on a mean luminance background) in order to map visual cortex regions that differ in angular and eccentricity representation. This allows for the delineation of primary visual cortex and extrastriate regions (Dumoulin & Wandell, 2008).

Secondly, finding the neural correlates of texture integration relies on a method described in Cornelissen et al. (2006). They introduced a method that allowed a so called signature plot of the strength of neuronal processes that resolve a contrast edge, caused by luminance differences between a disk and annulus, as function of cortical distance. As such, the signature is the amplitude of neural correlates, that resolve contrast edges, as function of cortical distance. Using texture-like stimuli, introduced in figure 1.4a, may allow to establish a signature of the integration processes as a function of cortical distance. Here, differences in the orientational



**Figure 1.4:** *Illusory Contour Induced by Orientational Properties.* (a) Differences in orientational properties exhort a percept of an illusory contour. The mean orientation of the Gabor patches between the disk and annulus differ substantially and therefore a percept of an illusory contour is visible. Integration processes, underlying the crowding effect, extract global features. In other words, the annulus and disk differ in their global feature characteristics. When fixation is directed to the green dot, the blue, red and yellow lines approximate the critical region, with eccentricity of the center at respectively  $4^\circ$ ,  $6^\circ$  and  $8^\circ$ . (b) The net-responses of the integration fields shows that the probability distributions differ around the illusory contour.

distribution of Gabor patches contained in the disk or annulus builds a percept of an illusory contour at the transition of the disk and annulus. Because integration fields responses are dissimilar near the illusory contour (integration fields lack consistency and withdraw a high sense of the illusory contour), we hypothesize that a signature of integration processes in response to the illusory contour can be obtained.

However, different mechanisms can endorse the percept of the illusory contour. Grigorescu et al. (2003) have shown that lateral inhibition, i.e. similar orientations in the vicinity of a cell's receptive field are inhibited, is a powerful strategy to detect edges and segment textures. Therefore, two conditions were implemented in the experimental paradigm to investigate the influence of lateral inhibition on the developed texture-like stimuli. Hence, in the lateral inhibition condition no noise is added to the orientational properties (standard deviation of the angle distribution of the Gabor patches  $VAR(\theta)_{disk} = VAR(\theta)_{annulus} = 0^\circ$ ). On the other hand, noise is added to investigate an illusory contour evoked by second-order angle differences attributed to processes in integration fields.

Note that, the visual system does not choose to use lateral inhibition as the strategy to detect edges. Instead, both feature integration and lateral inhibition provide visual cues to the perceived illusory contour. However, lateral inhibition shows only narrow local activity patterns and integration fields pool visual features in a region that extends to the size of the critical region in cortical space. Therefore, integration field responses are hypothesized to show a much broader response.

Lastly, the biologically plausible model for crowding was used to predict the signature of integration processes. Here, a virtual fMRI experiment is performed that followed the same experimental paradigm as the fMRI experiment. Thus, the orientational stimuli used in the fMRI signature experiment were modeled according to the crowding model, where the maximum amplitude of the net response of the integration fields were used as a measure for the profoundness of orientational information. These results provided a signature based on the crowding model and allowed the comparison between the signature obtained with fMRI and the biological plausible model for crowding.

### 1.3 Thesis Outline

This thesis aims to find neural correlates of integration fields in the early visual cortex using fMRI. If the reader of this thesis is not acquainted with brain activity imaging with fMRI, it is strongly advised to read the introduction of the fMRI technique in appendix A. Furthermore, some basic terminology in anatomical di-

rections and locations are outlined in appendix B. Chapter 2 provides background information in the structures and processes involved in human vision and especially crowding. Hence, the crowding effect and the theories about the underlying mechanisms of crowding are more thoroughly explained.

To find neural correlates of integration fields in visual cortical regions V1 and V2, the delineation of visual cortical regions was necessary. Visual field mapping techniques reveals the retinotopic organization of the visual cortex and enabled the delineation of cortical regions. The methods and results of the performed visual field mapping experiments are outlined in chapter 3. The methods and results used to find the neural correlates of integration fields are outlined in chapter 4. In chapter 5 the expected signature of the fMRI data is constructed using a biologically plausible model for crowding. Chapter 6 is a discussion of the obtained results. Furthermore, the comparison between the empirical and theoretical obtained data is outlined. The reason to outline the comparison between empirical and theoretical obtained data as late as in the discussion is to keep every chapter (explaining the performed experiments) independently readable.

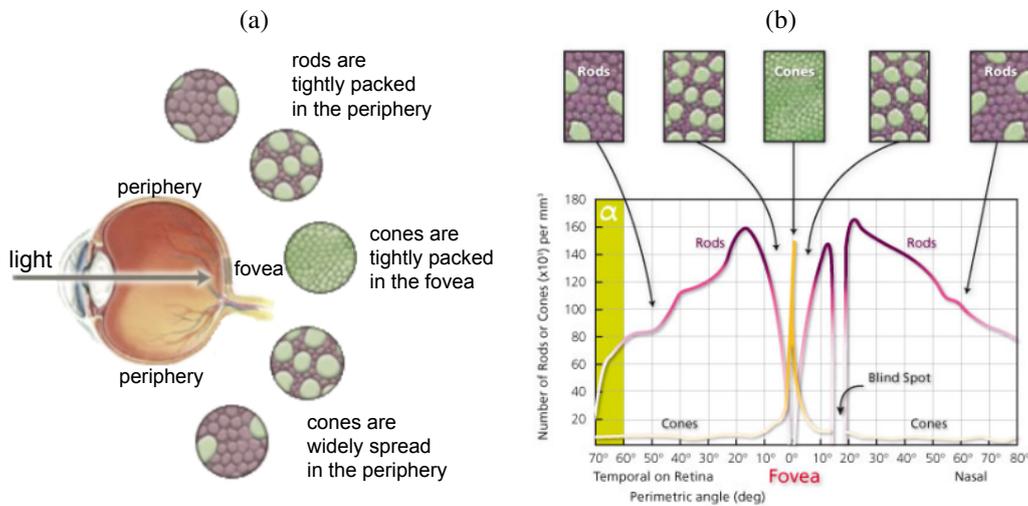
## Human Vision & Crowding

Human vision plays a prominent role in everyday life. Many cognitive tasks rely on visual processing, for instance, grabbing a beer from the fridge. Multiple objects need to be recognized to successfully grasp the beer without spilling the liquid. But human vision is more sophisticated and complex than sketched with this example. Humans are capable to extract meaning when a scene is presented for less than 100 milliseconds (Thorpe et al., 1996). Furthermore, even blurred versions of a presented scene provide enough information to understand meaning (Schyns & Oliva, 1994).

In the field of visual neuroscience, researchers are trying to unravel the underlying mechanisms of these fast and diverse visual-processing capabilities. This chapter provides an introduction into some important structures and processes that are involved in human vision. For a more detailed review of the structures and processes involved in human vision see Gazzaniga et al. (2002) and for a short review Grill-Spector & Malach (2004). Especially, the properties and theories about the underlying mechanisms of the crowding effect are thoroughly outlined.

### **2.1 Pre-Cortical**

The first stage of visual processing involves the eye (see figure 2.1a). The optics of the eye create an image of the visual world and project this image onto light sensitive tissue, i.e. the retina, located at the back of the eye. The retina consists of millions of photoreceptors that convert physical signals (light) into neuronal signals. There are two types of photoreceptors, rods and cones. Rods are very sensitive to light, but are unable to distinguish colors. Hence, rods play an important role in night vision. Cones are less sensitive to light than rods, but allow color



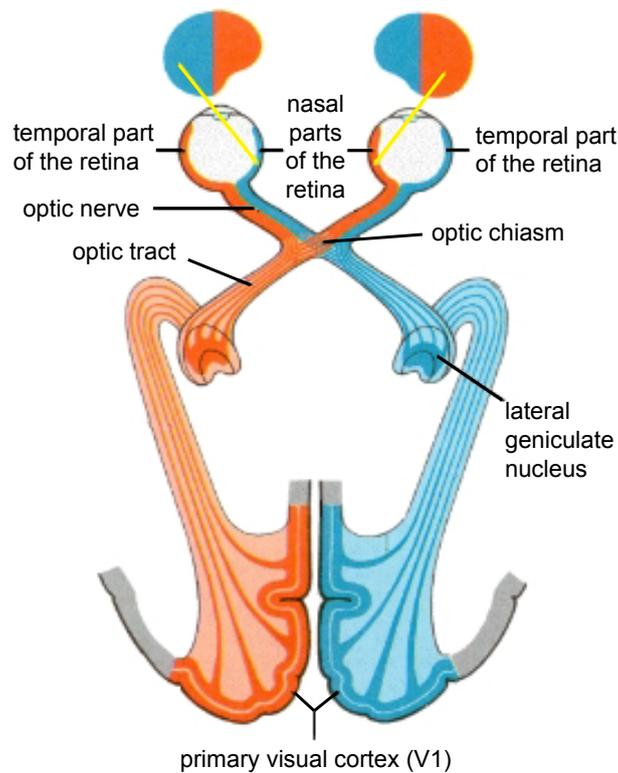
**Figure 2.1:** *The Eye. (a) The first stage of visual processing starts when light is projected through the optics of the eye at the back surface of the eye, i.e. the retina. The retina consists of millions of light-sensitive photoreceptors that convey the physical signal (light) into neuronal signals. There are two different types of photoreceptors. Cones are responsible for color vision and high acuity. On the other hand, rods are very sensitive to light and are responsible for especially night vision and motion perception. (b) In the fovea rods are absent but cones are highly dense. The periphery mainly consist of rods and their densities decreases towards the outerbounds of peripheral vision. (source graphics: <http://www.webexhibits.org>)*

vision and are responsible for high acuity (high resolution)<sup>1</sup>.

Densities of rods and cones in the retina are not consistent (Curcio et al., 1990), see figure 2.1b. The fovea is located at the center of the retina and only consists of cones. Hence, foveal vision is responsible for extracting small details and resolve color. Rods are absent in the fovea and are primarily found in the periphery of the retina. Thus, peripheral vision plays an important role in night vision. Densities of rods gradually decrease when moving towards the outer periphery, decreasing acuity even further.

After the visual world is conveyed into neuronal signals, the signals transverse through the optic nerves and is divided into two streams in the optic chiasm (see figure 2.2). The right visual field retinal branches (the temporal and nasal part of the retina of respectively the left and right eye) are projected to the contralateral

<sup>1</sup>In the retina photoreceptors have synapses to ganglion cells. Ganglion cells transport the neuronal signals further along the retinal branches. Many rods have synapses to one ganglion cell, whereas a single cone has a synapse to only one ganglion cell. Therefore, visual resolution is lower for rods in contrast to cones.

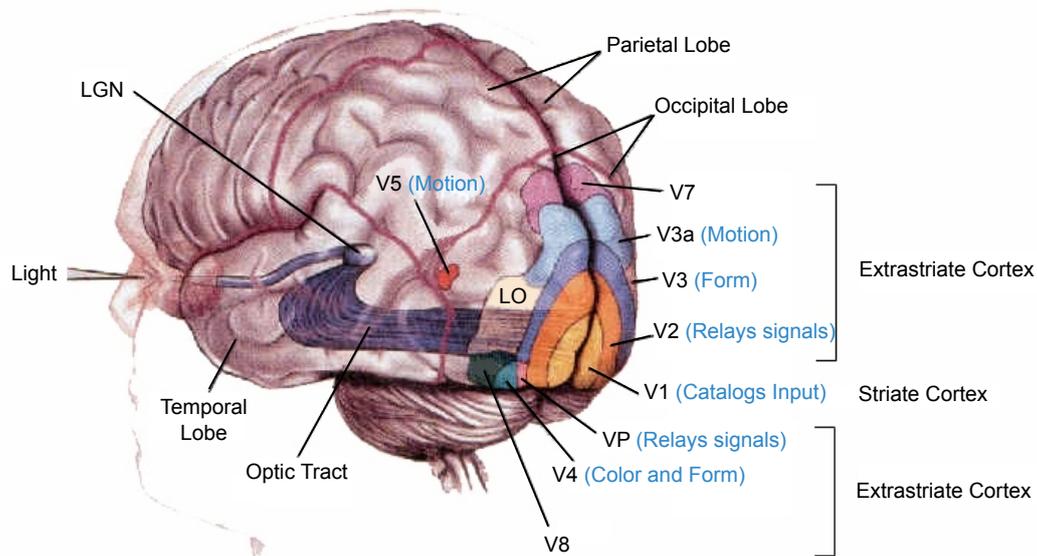


**Figure 2.2:** *The Retinal Branches.* The right and left visual field retinal branches are shown in respectively green and red. The nasal part of the left retina and the temporal part of the right retina receive projections of the left visual field. On the other hand, the right visual field is projected onto the nasal and temporal part of respectively the right and left retina. In the optic chiasm the temporal and nasal parts of both retina's are divided into a left and right visual field stream. The retinal branches transverse to the contralateral primary visual cortex through the optic tract. (source graphics: <http://colorado.edu>)

(left) hemisphere. The fibers of the right retinal branches, representing the left visual field, are projected to the contralateral (right) hemisphere. In other words, the left- and right visual field branches project onto the contralateral primary visual cortex (V1) located at the back of the brain.

## 2.2 Visual Cortex

There is a continuous spatial mapping, i.e. retinotopic organization, of the visual field onto visual cortical regions. Hence, adjacent points in the visual field are mapped onto adjacent neuronal populations in the visual cortex. Note that



**Figure 2.3:** *The Visual Cortex.* The visual cortex is located at the back of the brain in the occipital lobe. Nowadays, many different cortical regions devoted to visual processing are found. However, the location and cognitive functions are under constant debate. (source: <http://colorado.edu>)

the retinotopic organization of the visual field is always divided along the vertical meridian. That is, the left and right visual fields are respectively mapped onto the contralateral visual cortex. Furthermore, each cortical region consists of a representation of a complete map of the visual field. Therefore, cortical regions are determined by the boundaries of a complete map of the visual field. Decades of research have unraveled numerous visual cortical regions and their specialized functions, see figure 2.3 for a simplified layout of the visual cortex. However, the precise layout of visual cortical regions and their specialized function are under constant debate, especially extrastriate regions beyond V3. See Tootell et al. (1998) for a detailed review.

V1 has a very precise and orderly retinotopic organization. The reason for this is that cells in V1, compared to extrastriate cortical regions, have the smallest receptive fields. The retinotopic organization in extrastriate cortical regions is a bit more complex. For instance, the retinotopic organization of V2 is also divided along the horizontal meridian of the visual field. Thus, the lower and upper visual field quadrants are mapped onto respectively the dorsal and ventral part of V2. However, the retinotopic organization of the visual cortex remains continuous.

Another important property of V1, and most other extrastriate cortical regions, is that visual processing is more devoted to foveal vision. This discrepancy is due

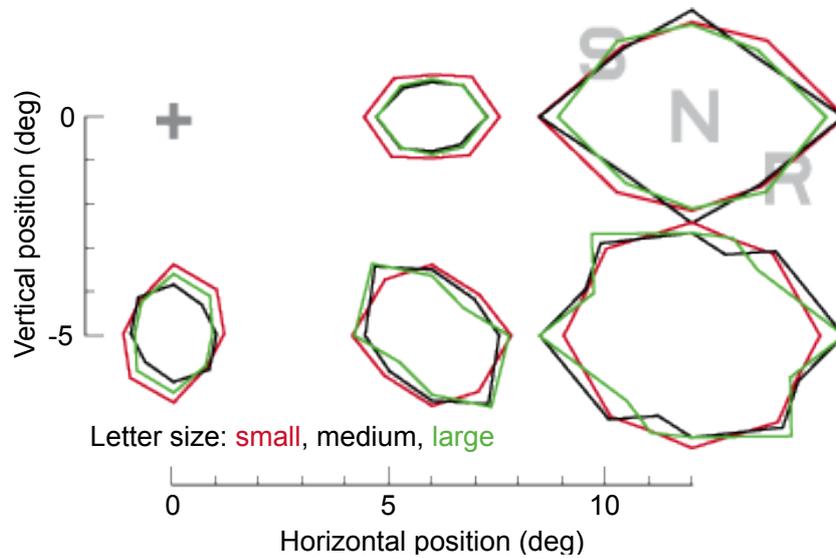
to cortical magnification, i.e. decreasing photoreceptor densities towards peripheral vision. In other words, the proportion of cortex devoted to foveal vision is significantly higher than in peripheral vision. The magnification factor is a logarithmic mapping of radial eccentricity and cortical distance in the visual cortex and varies between subjects and visual cortical regions. For instance, if cortical areas are more devoted to precise texture analysis, the cortical magnification factor is even higher and therefore the representation of peripheral vision is nearly absent.

Hubel & Wiesel (1977) used single cell recordings to investigate neuronal responses in the primary visual cortex. They found two most prominent cell types in the primary visual cortex and classified these as simple and complex cells. The receptive fields of simple cells have a central oval excitatory region which is surrounded by oval inhibitory regions. Therefore, simple cells are only responsive to bars or edges of a specific orientation in the area of the visual field. Complex cells receive input from many simple cells. As a result, it does not matter where in the receptive field of a complex cell a bar or edge is presented to elicit responses. Complex cells are found in the primary visual cortex (V1) and extrastriate cortical regions V2 and V3.

The findings of Hubel & Wiesel (1977), i.e. simple and complex cells are responsive to bars or edges, suggest that such cells are involved in elementary feature processing. The combination of elementary features allows a percept of an object composed of multiple elementary features (Pelli et al., 2009). As mentioned in the introduction, object recognition is severely impaired in peripheral vision. This could be attributed to lower acuity in peripheral vision which is reflected by decreasing photoreceptor densities and resolution towards the periphery or the crowding effect. In this thesis we are interested in the underlying mechanisms that endorse crowding. Especially, ‘hardwired’ visual feature integration in the early visual cortex is hypothesized to result in the crowding effect.

## 2.3 Crowding

The crowding effect entails the impairment of identification, in especially peripheral vision, of objects or features when surrounded by flankers and is independent of object size or type (van den Berg et al., 2007; Pelli & Tillman, 2008). Furthermore, crowding limits reading speed and is debated to be the underpinnings of disorders like amblyopia (Levi et al., 2007) and developmental dyslexia (Martelli et al., 2009). This section provides a short introduction of the crowding effect and the most prominent theories about the underlying mechanisms in the visual



**Figure 2.4:** *Critical Region of Crowding.* Crowding takes place in a critical region that is proportional to eccentricity. The vertical and horizontal axes represent position in the visual field relative to the fixation cross. The red, black and green circles are the boundaries of critical spacing for a letter identification task with respectively small, medium and large letter sizes. The critical region is unaffected by letter sizes. Therefore, crowding is independent of object size. (source: Pelli & Tillman (2008))

cortex. For a more detailed outline of the crowding effect see Levi (2008).

The first observations regarding crowding has been made by Korte (1923) in a letter identification task. For example, in upper part of figure 1.1, the target letter B on the left cannot, or merely, be identified when the centre of gaze is directed to the red fixation cross. The reason for this is that the flanker letters A surrounding the target letter B impairs recognition. Remarkably, Bouma (1970) showed that object recognition is unimpaired when center-to-center distance between object and flankers exceeds a critical spacing that is proportional to roughly  $0.5 \times$  eccentricity in the radial direction. This linear dependence of critical spacing and eccentricity is termed Bouma's law.

New insights reveal that critical spacing in the tangential direction is approximately  $0.1 \times$  eccentricity (Toet & Levi, 1992). The outerbounds of critical spacing result in a two-dimensional critical region wherein crowding takes place, see figure 2.4. Crowding is only prominent when objects fall within the critical region and therefore is scale-invariant and independent of acuity. For instance, letter size of target and flankers does not affect the spatial extent of the critical region (Pelli & Tillman, 2008), see different colored marked critical region in figure 2.4.

However, there are some discrepancies in the strength of the crowding effect. Crowding in the lower visual field is stronger than in the upper visual field (Intriligator & Cavanagh, 2001). Furthermore, objects that are projected closer in cortical space, although equally spaced in the visual field, show stronger crowding effects (Motter & Simoni, 2007; Liu et al., 2009). Hence, the crowding effect is less severe when objects are presented to the left and right of the vertical meridian because these are projected to the contralateral visual cortices. Additionally, this principle also results in the so called inward-outward anisotropy (Bouma, 1973; Petrov & Popple, 2007). Inward-outward anisotropy implies that an outward flanker (object presented at a higher eccentricity than the target) shows a stronger crowding effect on the target as opposed to an inward flanker. Although, the center-to-center distance in the visual field between inward- and outward flanker is equal, the outward flanker is closer in cortical space than the inward flanker. This asymmetry follows directly from cortical magnification.

Crowding is not only present in letter tasks, but also in for instance, orientation, size, saturation, and hue identification tasks (van den Berg et al., 2007), see examples in lower part of figure 1.1. Only similar features are crowded because, there is no crowding between tasks that are composed of different visual features. For instance, there is no crowding between first- and second-order composed letters (Chung et al., 2007). Despite the inability to accurately report flanked targets, Parkes et al. (2001) have shown that in an orientation discrimination task subjects readily report the average orientation of target and flankers. Hence, it seems that integration of similar visual features results in the crowding effect. Still, the underlying mechanisms of crowding are unknown. In the following section theories about the underlying mechanisms of crowding are explained.

### *2.3.1 Underlying Mechanisms of Crowding*

The detection of, for instance, a letter is explained through combining elementary feature detection mechanisms in the visual cortex, where elementary features are not labeled entities like ‘yellow’ or ‘square’ (Pelli et al., 2009). Feature detection channels are processing features like spatial frequency, bars or edges. In other words, when a feature detection neuron is presented with a stimulant inside its receptive field, detection of that specific feature takes place. Then, the elementary features are combined into a percept of an object. Thus, multiple features need to be detected and combined to acquire a true identification of the object or image. Crowding seems to affect the combining of features in an appropriate large area. Furthermore, the crowding effect cannot be attributed to masking or surround suppression (Pelli et al., 2004).

Masking occurs when an elementary feature detection unit is flanked, i.e. target and flanker influence the same feature detection unit, and therefore the target is not visible for the observer. In crowding conditions subjects report seeing a jumbled representation of the crowded signal. Hence, identification of elementary features is still possible. Moreover, crowding, in contrast to masking, is size-invariant and depends solely on critical spacing (Pelli et al., 2004). Another property of the visual system is surround suppression which is closely linked to lateral inhibition (Smith, 2006). Surround suppression is the decrease of neural activity when an optimal stimulus extended the neurons' receptive field (Hubel & Wiesel, 1977). This mechanism is dissimilar to the crowding effect, because inward-outward anisotropy does not exist in surround suppression (Petrov & Pople, 2007).

Nowadays, there are two prominent theories that explain the underlying mechanisms of crowding. Crowding is explained as the consequence of insufficient spatial resolution of attention (Intriligator & Cavanagh, 2001; Kouhsari & Rajimehr, 2005; Fang & He, 2008). For instance, Fang & He (2008) found attention enhancements of cortical responses when a target was surrounded with flankers. This enhancement did not exist when the target was non-attended. Furthermore, Pöder (2006) showed that when the number of distractors are increased, crowding is reduced. However, Dakin et al. (2009) showed that different neural mechanisms are involved for attention and compulsory averaging of crowded signals.

The most strongly supported explanation for crowding is 'hardwired' feature integration (Pelli et al., 2004). That is, neuronal populations in the visual cortex pool similar afferent visual features that fall within the critical region. Hence, integration of similar features within the critical region is a bottleneck for object recognition in especially peripheral vision. The spatial extent wherein neurons pool visual features, i.e. critical region, are addressed as 'integration fields' throughout this thesis. In other references, the region wherein integration takes place are termed 'combining field' (Pelli & Tillman, 2008) or 'association field' (Field et al., 1993).

Pelli (2008) argued that a fixed number of neuronal cells implement an integration field. The fixed cortical distance of the critical spacing in V1 is; given (exponential) cortical magnification, the linear Bouma's law is then log-transformed to a fixed cortical distance of 6 millimeter. Critical spacing in the tangential direction is much smaller. Here, the fixed cortical distance of integration fields were determined at approximately 1 millimeter in V1. In other areas, the fixed cortical distance of integration fields is different and depend on the cortical magnification of the investigated cortical region.

## Visual Field Mapping

Before the ability to investigate the signature of integration fields in the early visual cortex with functional Magnetic Resonance Imaging (fMRI, see Appendix A), the visual cortical regions in the parietal lobe should be delineated. The visual cortex is retinotopic organized, i.e. there is a continuous mapping of the visual field onto the visual cortex and each cortical region represents a complete map of the visual field (Wandell et al., 2007). Therefore, visual cortical regions are delineated by the boundaries of the retinotopic organization of a complete visual field representation. In other words, the delineation of the primary visual cortex (V1) and extrastriate visual cortical regions (V2/V3) makes use of the retinotopic organization of the visual cortex.

Visual field mapping techniques reveals the retinotopic organization of the visual cortex. This chapter provides an outline of the performed visual field mapping experiments and the delineation of the visual cortical regions V1/V2 and V3. The first section of this chapter provides an introduction of commonly used experimental paradigms and analysis methods of visual field mapping techniques. Hereafter, the materials and methods, i.e. fMRI acquisition and preprocessing, recorded subjects and stimulus description are outlined. Then, the results of the visual field mapping experiments are presented. That is, the resulting delineation of cortical regions (V1/V2/V3) in the visual cortex of the recorded subjects. For a review about this technique see Wandell (1999) and Wandell et al. (2007)

### **3.1 Visual Field Mapping Technique**

To find the retinotopic organization of the visual cortex stimuli are used that produce neural activity in the visual cortex. Checkerboard stimuli in shape of ex-

panding rings and rotating wedges are commonly used (Engel et al., 1994; Sereno et al., 1995; DeYoe et al., 1996; Engel et al., 1997; Dumoulin & Wandell, 2008). To attain large responses the checkerboards can be contrast reversing with a specific frequency or moving. The latter facilitates larger responses in V1 and other visual cortical regions and are now used extensively. Moreover, contrast reversing checkerboards are accompanied with lots of stress to the subjects due to the produced flickering of the fast changing contrast. The expanding ring stimulus (figure 3.1A) result in an eccentricity map of the visual field onto the visual cortex. Polar angle is mapped when using clockwise or counterclockwise rotating wedges (see figure 3.1B). An eccentricity and polar angle map together provides enough information to delineate the visual cortical regions.

Visual field mapping experiments mainly consist of multiple cycles of an expanding ring (from center to the border of the screen) and 360 degrees rotating wedge. Averaging fMRI responses of several runs of these stimuli increases the power of the signal and noise is reduced. The signal recorded in the fMRI scanner, in response to the presentation of contrast reversing checkerboards in shape of expanding rings or rotating wedges is conventionally analyzed with the traveling-wave method (Engel et al., 1994; Sereno et al., 1995; DeYoe et al., 1996; Engel et al., 1997). In short, the traveling-wave method implies a Fourier transform of the time series in every voxel<sup>1</sup>. In the frequency domain, the phase of the largest response in the time series is an estimation of the stimulus location in the visual field.

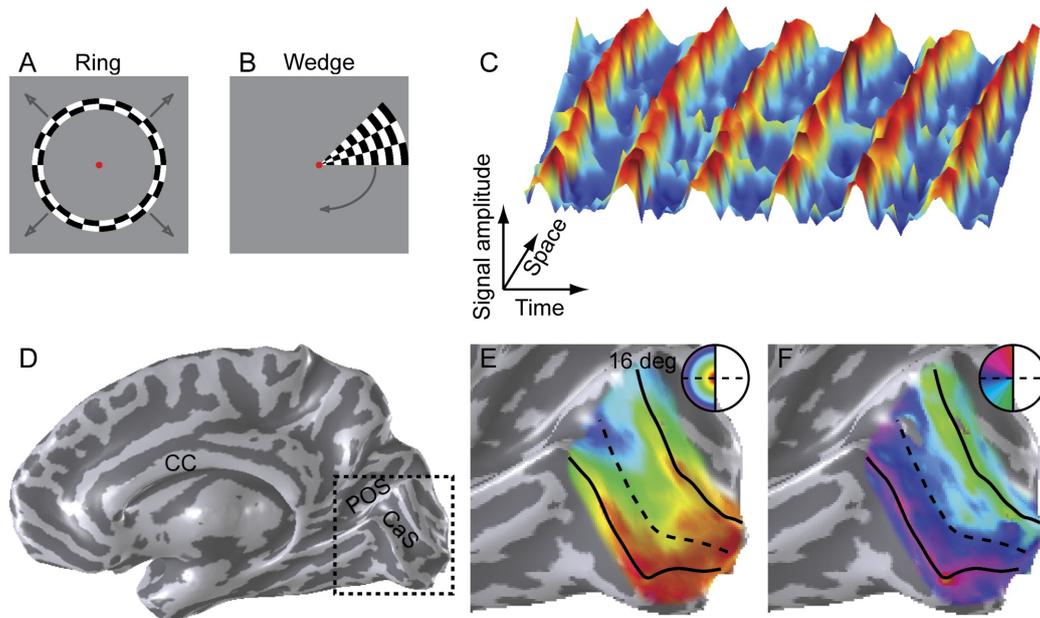
A more sophisticated method is the population receptive field (pRF) method (Dumoulin & Wandell, 2008). This method exploits the fact that neural activity extents more than a single location in the visual field. In other words, the extent of neural activity is determined by the population receptive field of a voxel. The pRF method provides a more accurate and reliable result. Furthermore, it provides a novel estimation in visual field mapping experiments; receptive field sizes of neuronal populations (voxels). In the following sections the traveling-wave method and the pRF method are explained.

### 3.1.1 *Traveling-wave method*

The most straightforward method to analyze retinotopic maps is the traveling-wave method (or phase-encoded retinotopic mapping). The traveling-wave method is the analyses of the largest response of the time series in every voxel in response to moving visual stimuli. By expanding and rotating respectively the ring and

---

<sup>1</sup>A voxel is the smallest volume fMRI measurement and represents a three-dimensional location in the brain.



**Figure 3.1:** *Traveling-Wave Method.* The widely used stimuli in visual field mapping are (A) the expanding ring for eccentricity mapping and (B) the rotating wedge for polar angle mapping. (C) These stimuli produce a traveling wave in the visual cortex due to the retinotopic organization. There are 6 traveling waves (red indicates the highest response to the stimuli). Hence, the experiment consisted of six cycles of a expanding ring from center to the border of the screen or 360 degrees rotating wedge. (D) An inflated mesh is used to visualize the responses. The traveling waves are calculated for the voxels in the visual cortex, important structures that are labeled are the corpus callosum (CC), the parietal-occipital sulcus (POS) and calcarine sulcus (CaS). The resulting phase maps are visualized on an expanded view of the inflated mesh and represent respectively the eccentricity (E) and polar angle (F) map. The solid black line is the boundary of the primary visual cortex (V1). (source: Wandell et al. (2007))

wedge stimuli adjacent cortical regions progressively shows activity patterns in the time courses due to the retinotopic organization of the visual cortex. Therefore, the stimulus creates a continuous traveling wave of neuronal activity in space (see largest response in space in figure 3.1C). After a complete cycle of the stimulus (ring expands from center to edge of screen or 360 degrees rotating wedge), another presentation of a cycle will produce a similar traveling wave of neural activity.

The traveling-wave method implies that the time series in every voxel in the visual cortex is analyzed with a Fourier transform. This manipulation allows a description of the time series of every voxel in the frequency domain. The phase of the stimulus modulation frequency in Fourier space of the time series of a voxel unfolds the correspondence between the location of the stimulus in the visual field and the visual cortex (the frequency of the largest response in a voxel is similar to the frequency of the presented stimulus modulation). In figure 3.1E&F the phase for every voxel in response to respectively the expanding ring and rotating wedge stimulus frequency is projected onto an inflated mesh. The eccentricity and polar angle map reveals the retinotopic organization of the visual cortex.

### 3.1.2 *Population receptive field method*

Dumoulin & Wandell (2008) describe a more sophisticated model-based method which is more accurate and reliable than the traveling-wave method. This method exploits the fact that neural activity extends more than a single location in the visual field. Hence, the extent of neural activity is determined by the population receptive field of a voxel (pRF). See figure 3.2 for a graphical representation of the steps involved in the analysis of visual field mapping experiments as depicted by the pRF method.

First, the pRF for every voxel is independently modeled as a Gaussian envelope of neuronal population activation. Then, a prediction of the time series for every voxel is calculated according to the time course of the stimulus aperture and the neural responses of the modeled pRFs. The stimulus aperture is a time series of the stimulus positions for every scan, where all visual field mapping experiments and the corresponding fMRI data are concatenated (rotating wedge and expanding ring). The expected time series is convolved with the hemodynamic response function (HRF) to compute the prediction of the BOLD-response and this is fitted with the fMRI data. Hereafter, parameters of the pRF are altered and a new prediction is computed. The best fit between the predicted fMRI signal and the observation result in an accurate and reliable retinotopic organization of the visual cortex, where the center of the modeled pRF corresponds with the location of the

stimuli in the visual field, i.e. eccentricity and polar angle. Furthermore, a novel parameter in visual field mapping techniques is measured with the pRF method; the size of the population receptive field of a voxel.

The pRF method does not rely on a periodic, i.e. phase-encoded, stimulus presentation. Therefore, other stimuli descriptions can be developed. For instance, Dumoulin & Wandell (2008) introduced a moving bars experiment for better foveal mapping. Furthermore, blanks (no stimulus presentation) can be inserted in the stimulus aperture to reduce adaptation effects and disturbances of constant stimuli presentation. The exact methods used to delineate the visual cortical regions is outlined in the following section.

## 3.2 Materials and Methods

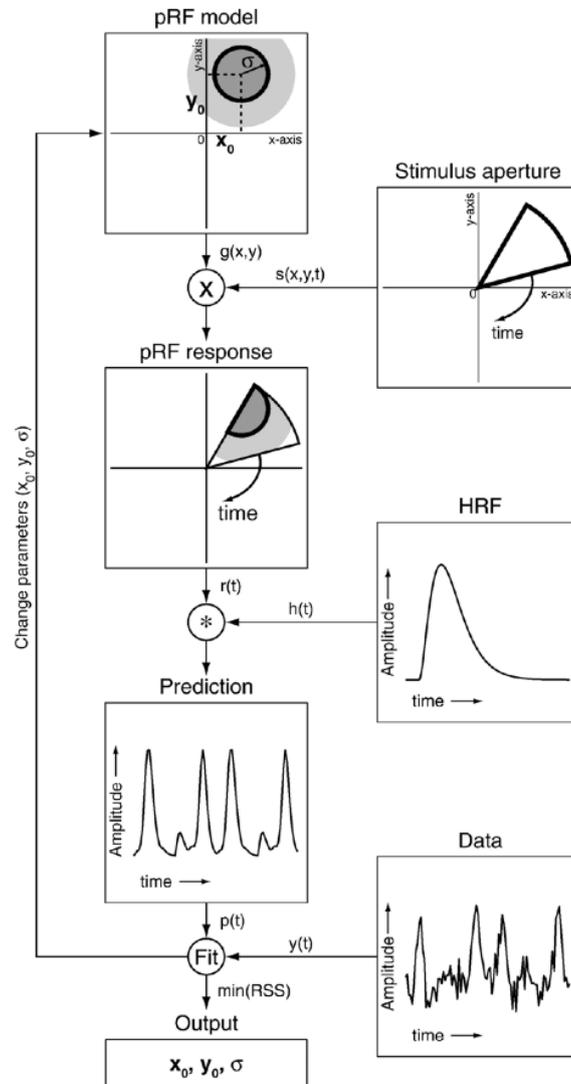
The main toolbox used for the preprocessing and analysis of the fMRI data is the *vistasoft* (f)MRI analysis toolbox developed in Matlab (MathWorks, Natick, MA) at Stanford University (available for free at <http://white.stanford.edu/>). This toolbox provides the *mrVista* program and supplementary (f)MRI data processing tools. In the *mrVista* program, all (f)MRI data were imported, visualized and analyzed. If reader is not familiar with the fMRI technique see Appendix A for a short introduction.

### 3.2.1 Subjects

In the field of visual field mapping, the usual approach is to accurately map visual fields in a relatively small number of subjects. Subjects are analyzed individually and their data is also presented individually. The experiments were performed on two subjects, SF and KH, both with normal vision. Both subjects are right-handed and are aged respectively 25 and 27. The recruited subjects are colleagues with previous experience as a subject in fMRI experiments, particularly visual field mapping experiments, and are readily available. Previous experience in fMRI experiments reduces motion confounds.

### 3.2.2 Magnetic Resonance Imaging

The (f)MRI data is acquired with a 3T Philips MRI-scanner in combination with a 8-channel head coil. For every subject a detailed T1-weighted contrast image, i.e. a full anatomy, of the whole brain is recorded and provides a detailed hydrogen



**Figure 3.2:** Population Receptive Field Method. Approach of the pRF method. For every voxel the parameters and models the center location in the visual field of a voxel and represents the size of the receptive field. A prediction of the time series in response to the stimulus aperture is computed. Then, the convolution of the time series prediction and the HRF result in the expected observation of the fMRI data. The parameters are changed and a new prediction is computed. The best fit to the fMRI data results in an accurate and reliable retinotopic organization of the visual cortex. (source: Dumoulin & Wandell (2008))

contrast image of the brain. The dimension of the voxels in the full anatomy is  $2.33 \times 2.33 \times 3 \text{ mm}^3$ .

During the presentation of visual field mapping experiments we are interested in the BOLD-responses in the visual cortex. The BOLD-responses are recorded in so called functional scans which is a partial brain recording. The partial brain recording is a  $96 \times 96 \times 24$  voxel grid with a specified position and direction. The setup of the sagittal, axial and transverse position and direction of the partial recording is determined by the experimenter at the beginning of a scan-session of a subject. In this research the main focus lies in early visual processing in the visual cortex. Hence, the recording position of the functional scans minimally included the visual cortex and is oriented parallel to the Calcarine sulcus (where the primary visual cortex is located). The size of the voxels is small enough to investigate the size of the integration fields. Integration fields implies pooling in V1 over a fixed cortical distance of 6 mm with regard to eccentricity (Pelli, 2008). The BOLD-responses of every voxel in the functional data need to be mapped onto the anatomy of the brain. To allow a robust alignment of the partial brain scan and the full anatomy, a low detailed T1-weighted inplane anatomy is recorded on the same position and has the same dimension as the functional scans.

Another important parameter is repetition time (TR) which determines the duration of one scan (measurements of relaxation periods for every voxel is not instant). A larger TR allows higher resolutions but prohibit 'fast' alternating stimuli. It is common when measuring the BOLD-response a  $\text{TR} \geq 1500$  milliseconds is used. A partial functional brain scan focusing on the visual (parietal) cortex allows a TR of 1.5 seconds in combination with an inplane resolution of  $2.0 \times 2.0 \times 2.0 \text{ mm}^3$ .

### 3.2.3 Anatomical preprocessing

All data is recorded in the radiological convention or Left-Anterior-Superior orientation (LAS) and are converted to the neurological convention, i.e. Right-Anterior-Superior (RAS) orientation. Consequently, only the right-left orientation is swapped. The orientation of the neurological orientation is the same as the Talairach template coordinate system. The Talairach template coordinate system uses standard anatomical landmarks that enable the comparison between subjects (Talairach & Tournoux, 1988).

The preprocessing of the full anatomy consists of two steps. Firstly, a realignment to anterior-commissure-posterior-commissure-space (AC-PC space) was performed. For this realignment, the Ac and PC coordinates in the anatomy of each subject was determined. The AC is the origin of the Talairach coordinate system

(position 0,0,0). It is a small spot of white matter anterior to the thalamus and connects the two hemispheres. This is different from the corpus callosum which is a much larger connection between the hemispheres. A reference point to the AC point is the PC point which is a few centimeters posterior to the AC and looks like a small fiber tract connecting the cerebellum with the thalamus. Secondly, voxel-sizes are re-sliced to 1 millimeter<sup>3</sup> isotropic voxels (uniformity in all directions).

### 3.2.4 *Functional preprocessing*

The first 12 seconds of each functional scan session is discarded from further analyses to minimize transient magnetic saturation effects in the time series. To account for head movement artifacts, motion correction between sessions was performed using the rigid-body technique (Nestares & Heeger, 2000). Hence, each functional recording session is motion corrected to a reference functional scan. The reference scan for the between-session motion compensation is the scan before to the inplane recording (there are techniques that use the inplane anatomy for motion correction but this imposes computational and accuracy issues). Within-session motion correction is not performed, because the functional scans of a single experiment are of a small duration. Hence, there is no significant influence on the quality of the results and therefore within-session motion compensation is redundant.

After preprocessing anatomical and functional data, the functional scans were aligned to the full anatomy. The inplane anatomy is recorded on the same position and dimension grid as the functional scans. Therefore, transformations of the inplane to the full anatomy were used to align the functionals to the full anatomy. The alignment of the inplane- to the full anatomy is performed using mutual information (Collignon et al., 1995). Lastly, functional data that was recorded during similar stimulus presentations were averaged to reduce signal-to-noise ratio.

### 3.2.5 *Gray and white matter segmentation*

Visual processing and other cognitive processing functions are represented in grey matter tissue of the brain and is mainly distributed along the surface of the brain. On the other hand, white matter only conveys information between structures. Hence, only the BOLD-responses of voxels that is classified as grey matter were analyzed. The most straightforward method to classify voxels as grey matter is to grow grey nodes to a white matter segmentation (classifying the white matter voxels). An automatic white matter segmentation was computed using FSL (Smith et al., 2004; Woolrich et al., 2009). This method results in a rough segmentation

of sub-cortical brain structures and white tissue. Furthermore, cerebrospinal fluid (CSF) is segmented to prohibit grey matter classification as CSF. The anatomical images were noisy, therefore automatic segmentation did not yield a satisfactory result and manual segmentation was necessary.

ITK-Snap is a freeware program with many tools to accurately make manual adjustments to a segmentation (Yushkevich et al., 2006). Furthermore, classifying irrelevant structures as CSF will ensure that no other tissues or structures are classified as grey matter. For instance, classifying a border between the left and right hemisphere ensures grey matter is not grown in the other hemisphere. A useful extra feature of ITK-Snap is a three-dimensional image reconstruction of the segmentation. This reconstruction allows a three-dimensional view of the segmentation and allows easier detection of segmentation errors. After the white matter is segmented, three layers of gray matter are grown to the white matter segmentation, CSF excluded, using ITK-Gray (part of the ITK-Snap project). The resulting grey nodes are imported into mrVista.

### 3.2.6 *Visualizing the visual cortex*

The imported grey nodes in mrVista restrict the functional data to grey matter. However, visualization on two-dimensional anatomical images prohibits a clear view of the precise outlay of BOLD-responses. The reason for this is the complex surface properties of the human brain. To visualize BOLD-responses, the white matter segmentation can be visualized as a mesh. A mesh is a three-dimensional reconstruction of the segmentation and data can be projected onto the surface of this mesh. Furthermore, this mesh can be inflated and smoothed to enhance the clarity and visibility of stimulus responses.

Another extensively used method is creating a flattened two-dimensional view of the grey matter nodes. This procedure needs a starting point that determines the center of the flattened cortex. From this point the cortex will be flattened (Teo et al., 1997; Wandell et al., 2000). In other words, the distance of the voxels in three-dimensional space to the starting point is approximately represented by the radial distance of the two-dimensional surface of the flattened cortex. Hereafter, responses are projected onto the flattened surface.

### 3.2.7 *Stimulus presentation*

Subjects viewed the stimuli through a back-projection screen placed at the end of the magnet core. Through an adjustable mirror placed on the head-coil, the subjects could see the back-projection screen. The size of the beamer projection onto

the back-projection screen was  $44 \times 33$  cm. The viewing distance, i.e. total distance from the eyes to the screen, was 75 cm. Hence, the corresponding maximum visual angle (eccentricity) of the stimulus presentation is  $32.7 \times 24.8^\circ$ . Resolution of the beamer was set to  $800 \times 600$  pixels.

The experiments conducted in this research depend on the center of gaze (because of the retinotopic organization of the visual cortex). During the presentation of stimuli a fixation point is always presented on the center of the screen. Furthermore, subject are asked to report, by pressing a button, every change of color of the fixation point to make sure that the subject is fully focused on the fixation point.

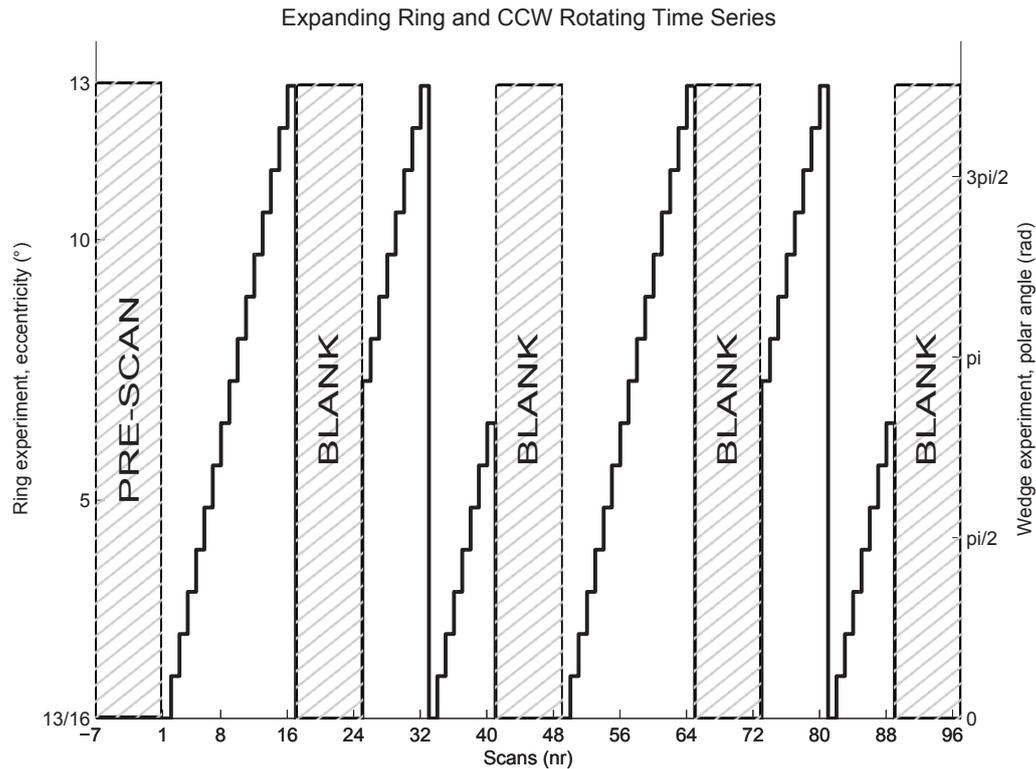
### 3.2.8 Stimulus description

The visual stimuli used in the visual field mapping are created using Matlab (MathWorks, Natick, MA) and supplementary tools provided in the Psychtoolbox (Brainard, 1997; Pelli, 1997) and vistasoft (f)MRI analysis toolbox developed in Matlab (MathWorks, Natick, MA) at Stanford University (available for free at <http://white.stanford.edu/>). The vistasoft package provides standardized visual field mapping experiments. Furthermore, the stimuli are developed for analyzing the responses with the pRF method.

The stimuli are presented on a mean luminance background and consisted of checkerboard contrast patterns (black and white) radially drifting from or away from fixation point with 8 frames per second. See figure 3.3 for the experiment paradigm of a single run of the rotating wedge and expanding ring stimuli. The wedge extended  $45^\circ$  of visual angle and the ring had a width of  $13/16^\circ$ . Each run consist of 6 cycles where the period of a complete cycle of the stimuli (full rotation and full expansion of respectively the wedge and ring) has a duration of 16 scans. The stimulus advances in eccentricity or polar angle in synchrony with the acquisition rate that is determined by the TR. The resulting run length is  $16 \times 6 \times 1,5 = 144$  seconds.

Before each run a 12 seconds period of a mean luminance screen is presented to take the scanning distortions at the start of the scanning procedure into account (magnetic transient effects). The stimuli are developed for the pRF method and therefore mean luminance screens are inserted with a frequency of 4 and a duration of 8 scans (Dumoulin & Wandell, 2008). The first mean luminance screen is inserted at the end of the first cycle.

An additional moving bars experiment is performed to provide a more accurate mapping of foveal parts of the visual field. The moving bars experiment consists of moving checkerboard patterns in shape of a bar and originates from each corner



**Figure 3.3:** *Visual Field Mapping Experiment Paradigm.* Visual field mapping experiment paradigm of the expanding ring and counterclockwise rotating wedge stimulus. The figure shows the progression of the stimuli in scans. The paradigm exist of 6 complete cycles, with a duration of 16 scans per cycle, and 4 mean luminance blocks with a duration of 8 scans. The pre-scan duration is 8 scans (12 seconds) and is removed from further analysis. One cycle of the expanding ring, left y-axis, starts from the center of the screen and progresses to the outer border of the screen. One cycle of the counterclockwise rotating wedge, right y-axis, consists of a 0 to  $2\pi$  radial position.

or cardinal point and transverses to the opposite position of the screen. A run of the moving bar experiment consist of 1 cycle where the origin of the bar-stimulus is successively north, west, south, east, northwest, southwest, southeast and northeast. Hence, the origin of the bar-stimulus is every corner and cardinal point of the screen and crosses the screen 8 times. Each crossing has a duration of 16 scans and therefore is divided into 16 equally spaced steps from origin and the opposite of the screen.

Mean luminance screens are inserted in the whole second-half of each crossing. Hence, the perceived bar stimulus only transverses from the corner or cardinal point to the middle of the screen. The total duration of the moving bars experiment is  $16 \times 8 \times 1.5 = 192$  seconds. Furthermore, a pre-scan duration of 12 seconds is inserted. Note that the moving bars experiment is not phase-encoded. In other words, the period of the bars stimulus is never repeated.

The subject and hemisphere dependent shape of the Hemodynamic Response Function (HRF) influences the results of the pRF sizes. Therefore, the HRF is analyzed for each hemifield of the subjects using a full-field stimulus. A cycle of the full-field stimulus is a screen filling contrast moving checkerboard for duration of 2 TRs and a blank screen with a duration of 30 seconds. Thus, each cycle has a total duration of 33 seconds and a single run contains 6 cycles. As a result of the screen filling stimulus, two pulses of full activity in the visual cortex are recorded. This corresponds with two linear additive pulses of the HRF. The time series in V1 at eccentricities between 5 and 9 degrees are extracted from the full-field experiment and are fitted with a two-gamma model of the HRF, where the HRF is modeled as a space-time invariant convolution kernel (Logothetis & Wandell, 2004). We use the two-gamma model instead of the simple canonical HRF that is described by the Boynton model (Boynton et al., 1996). The two-gamma model is a better approximation of the HRF because in the Boynton model the undershoot before and after the peak response are not modeled. See 3.2 for the description of all performed experiments.

### 3.3 Results

The realignment of the functional data onto the anatomical recording of the subjects yielded a satisfactory result. In the left part of figure 3.4, the reconstructed position and direction of the partial brain scan for subject KH is shown. This reconstructed image follows from the transformation of the inplane-anatomy to the full-anatomy recording. Functional data is realigned to the full-anatomy using the same transformations. Furthermore, functional data is restricted to gray matter

Experiment	stimulus period (sec)	NR of cycles	NR of runs	Total duration (min)
Session 1: E/C rings	24	6	8	27.2
Session 2: CW/CCW wedge	24	6	8	27.2
Session 3: 8 bars	210	1	4	18
Session 4: full-field	33	6	4	13.2

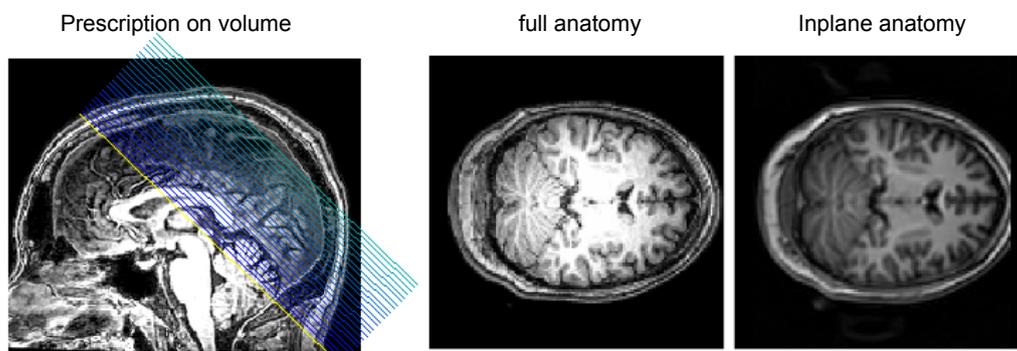
**Table 3.2:** *Visual Field Mapping Experiments.*

using the segmentation of white matter and CSF, see figure 3.5 for a axial, sagittal and coronal slice of the segmentation.

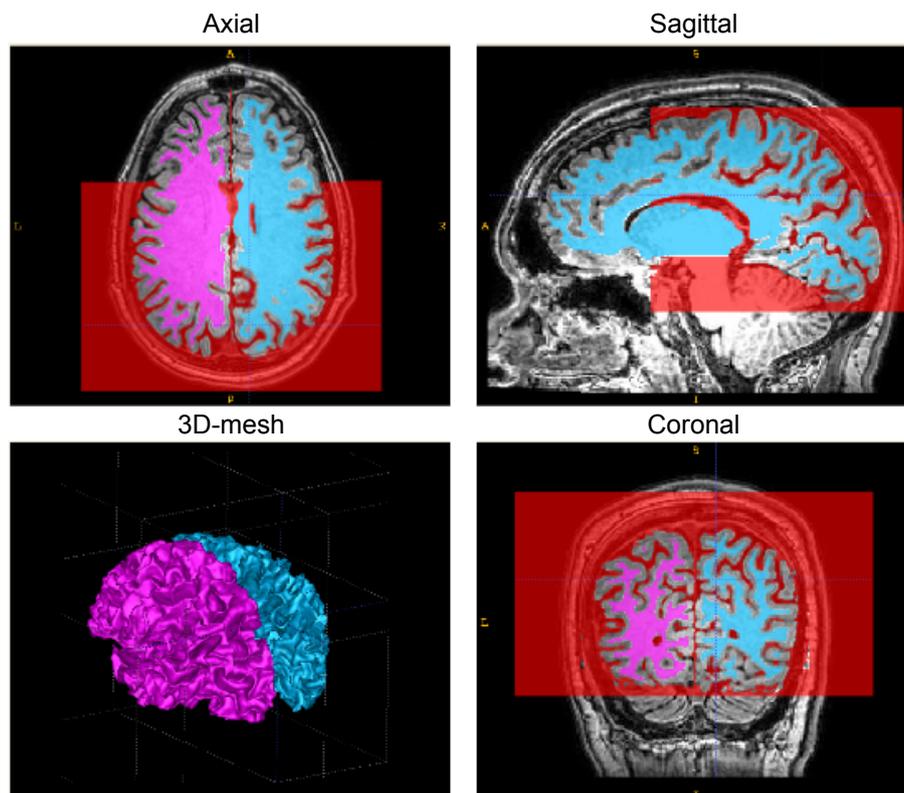
The analysis of the visual field mapping experiments reveals the retinotopic organization of the visual cortex. An eccentricity and polar angle map provides enough information to allow the delineation of the primary and extrastriate cortical regions. The two different methods explained extensively are the traveling-wave and the population receptive field model (pRF) method. The latter provides a more accurate and reliable result and is therefore used extensively for extracting the retinotopic organization of each subject.

However, first a traveling wave analysis is performed that allowed the delineation of V1 at eccentricities 5 to 9 degrees. Here, the time series in response to the full-field experiment is used to analyze the subject dependent HRF. In figure 3.6 the time series of the full-field experiment in V1 at eccentricities between 5 and 9 degrees in each hemifield is shown for subject KH. The filled circles represent the average BOLD-signal of the data and the solid line is a two-gamma fit to the data. Because the full field stimulus is presented for two TRs, the data and two-gamma model represent two impulses of the HRF. Hence, the dotted line is the approximation of a single HRF pulse.

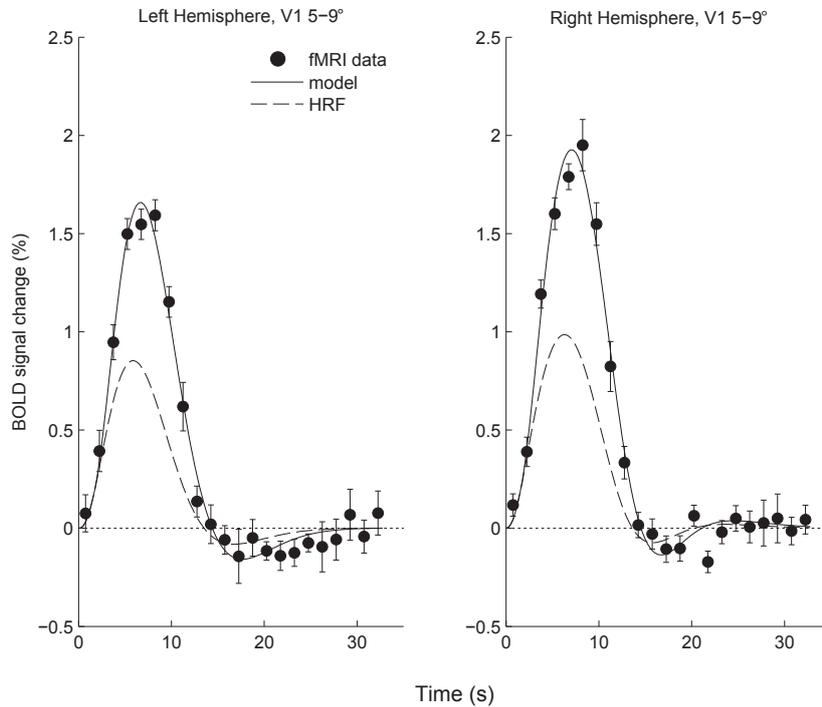
Then, the pRF method in combination with the subjects HRF is performed on the functional data of the visual field mapping experiments. As a result, see figure 3.7A&B for a polar and eccentricity map in the left visual cortex of subject KH. The retinotopic data is visualized on an inflated mesh and flattened cortex. The center of the flattened views originates approximately from the foveal part of V1.



**Figure 3.4:** *Alignment.* The result of the alignment of the inplane anatomy to the full anatomy. In the left side of the figure, the blue lines resemble the reconstructed recording position of the inplane anatomy opposed to the full anatomy. The two pictures on the right side show the yellow slice of the full and inplane anatomy.



**Figure 3.5:** *Segmentation.* Axial, sagittal and coronal slice of a white matter and CSF segmentation of subject KH. Purple and blue respectively indicate segmentation of white matter in the left and right hemisphere. CSF is segmented with red. At the lower left of the figure the mesh, i.e. three-dimensional reconstruction of the white matter segmentation, is shown.



**Figure 3.6:** *HRF Estimation of Subject KH. The filled circles is the average time serie in response to the full-field stimulus in respectively the right and left primary visual cortex (V1) and is restricted to voxels that represent 5 to 9 degrees of visual angle. The solid line is a two-gamma fit to the data. the dotted line is the HRF.*

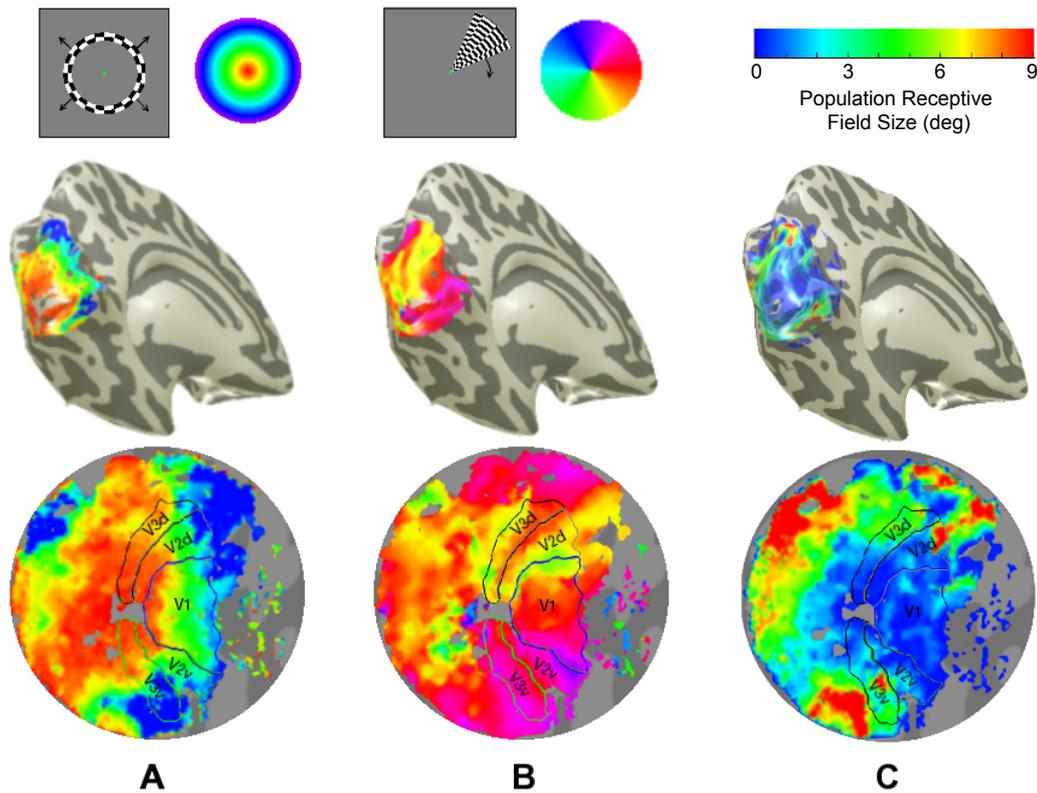
The delineation of the visual cortical regions was done by hand on the flattened cortex. We are only interested in the responses in the primary visual cortex (V1) and extrastriate cortical region V2. Therefore, the borders of higher extrastriate cortical regions were not determined.

The eccentricity map in figure 3.7A shows the radial mapping of the visual field onto the flattened cortex. The legend of the eccentricity map is the center, in the visual field, of the best pRF model fit. Note that the visual cortex extends to an eccentricity beyond the mapped eccentricity. However, the delineation of more peripheral parts is impossible due to the maximum size of the display inside the MRI scanner. In other words, the peripheral border of the visual cortical regions extent to the maximal eccentricity of the experimental setup.

In figure 3.7B, the polar angle map of the visual cortex in the left hemisphere is shown. The cortical regions have a complete representation of the contralateral right visual field. The representation of the vertical meridian is the border of V1. After V1, the dorsal and ventral part of the extrastriate region V2 represents

respectively the lower and upper quarterfield. Hence, the border of the dorsal and ventral part of V2 is the vertical meridian, i.e. the borders of V1, and the horizontal meridian. For V3 the opposite holds, the lower and upper quarterfields of the right visual field are represented in respectively the ventral and dorsal part of V3. Additionally, the pRF sizes of voxels were computed and are shown in figure 3.7C.

The data shows similar results as obtained by Dumoulin & Wandell (2008). However, retinotopic maps can be more accurately obtained when projected onto an inflated mesh instead of a flattened cortex. The disadvantages of a flattened cortex is that there is an extra analysis step involved. But, more importantly distortions of cortical distances are known to occur in the flattening procedure.



**Figure 3.7:** *Retinotopic Organization of Subject KH. Position estimates of the pRF method of subject KH. The upper part of A and B shows the stimulus aperture and the legend of the phase and the corresponding location in the visual field. The retinotopic maps of the left visual cortex representing the right visual field are projected onto an inflated mesh and flattened cortex. The retinotopic map consists of a (A) eccentricity and (B) polar angle map. (C) Additionally, the population receptive field sizes of the voxels were determined and projected onto an inflated mesh an flattened cortex. The delineation of the visual cortical regions V1, V2 and V3 are shown for every condition in the flattened cortex and was done by hand.*

# Chapter 4

## Signature Integration Fields: an fMRI approach

This chapter is an outline of the techniques, methods and results of finding neural correlates of integration fields in the early visual cortex using fMRI. In the previous chapter the results of the visual field mapping experiments are outlined. The delineation of visual cortical regions allows the search for the neural correlates of integration fields in the early visual cortex (V1 and V2). We hypothesize that these neural correlates can be found by investigating the so called signature of integration processes. Finding this signature relies on a method introduced in Cornelissen et al. (2006). Here, a signature of neural responses to a contrast edges is investigated. As such, the signature resembles the strength of neural activity as function of cortical distance to the perceived contrast edge in the early visual cortex. Using a similar strategy, we developed texture-like stimuli to investigate neural correlates of integration processes ascribed to integration fields. Integration fields are hypothesized to be involved in 'hardwired' feature integration and this integration is closely linked to the crowding effect.

The first section of this chapter provides an introduction of the technique used to reveal the signature of neural processes. Hereafter, the methods, i.e. the developed texture-like stimuli and performed experiments are explained. Lastly, the analyzed signatures of integration processes in response to the developed texture-like stimuli that are hypothesized to resemble integration processes are presented.

## 4.1 Signature of Neural Processes

The signature of neural responses in the visual cortex using fMRI can be determined by a method similar to that described in Cornelissen et al. (2006). Cornelissen et al. (2006) developed a stimulus description and experimental paradigm that enabled a signature of neuronal processes in response to a contrast edge. The signature of neuronal processes resolving contrast edges was defined by the strength of neural response to the contrast edge as function of cortical distance to the perceived edge.

The technique to reveal a signature of edge resolving neural processes is constructed as follows. Subjects are presented with a mean luminance screen where a central disk is defined with a radius of half the maximum radius of the screen. The task for the subjects is to keep their center of gaze directed to a fixation point presented in the center of the screen. Through modulating the luminance around the mean luminance of the background of either the disk or surround with a sinusoid a contrast edge is perceived at the transition between the disk and surround. Perceptually a similar observation as the modulation sinusoid is prominent, i.e. the strength of the perceived contrast edge follows the modulation sinusoid. Moreover, neural processes resolving contrast edges exhibit a similar modulation as the luminance modulation. Then, a fast Fourier Transform is used to analyze the amplitude of the BOLD modulation in every voxel in the visual cortex.

To reveal the signature of neural responses to the contrast edge as function of cortical distance, the transition of the disk and annulus is mapped onto the visual cortex using a center-surround localizer. The center-surround localizer is a contrast moving checkerboard in shape of the disk and annulus and are sequentially presented. Hence, voxels that respond to either the disk or annulus are revealed in the visual cortex and the transition is the representation of the contrast edge in cortical space. Now, the amplitude of neural processes in response to the contrast edge can be plotted as function of cortical distance to the transition of the disk and annulus in the visual cortex. Note that the cortical distance is an approximation of the distance in the visual field. The reason for this is the retinotopic organization of the visual cortex and the fixed eccentricity of the contrast edge, i.e. transition between disk and surround is at a fixed eccentricity from fixation.

### 4.1.1 Signature integration fields

Finding neural correlates of integration processes relies on the same technique. However, in this study we are interested in integration processes attributed to integration fields. Therefore, texture-like stimuli were developed that evoke integra-

tion processes. In the introduction, i.e. figure 1.4(a), an example of the developed texture-like stimulus is shown. The stimulus consists of many oriented Gabor patches. Differences in second-order<sup>1</sup> orientational properties between the central disk and surrounding annulus induces the percept of an illusory contour at the transition. Differences of the net responses of integration fields are hypothesized to produce the perceived illusory contour.

Through modulating the difference of the orientational properties of the Gabor patches contained in the disk and annulus integration processes are exhibited. Hence, we hypothesize that the modulation of orientational differences between the disk and annulus evokes 'hardwired' feature integration. In other words, neural processes that involve feature integration show similar modulation as the stimulus modulation. Thus, this method allows an amplitude signature plot as function of cortical distance to the illusory contour. Hence, similar integration processes with similar spatial extents are averaged.

## 4.2 Materials and Methods

The main toolbox used for the preprocessing and analysis of the fMRI data is the *vistasoft* (f)MRI analysis toolbox developed in Matlab (MathWorks, Natick, MA) at Stanford University (available for free at <http://white.stanford.edu/>). This toolbox provides the *mrVista* program and supplementary fMRI data processing tools. In the *mrVista* program, acquired (f)MRI data were imported, visualized and analyzed.

### 4.2.1 Acquisition and preprocessing

The visual field mapping experiments described in chapter 3 provides a framework to investigate the signature of texture integration. This framework is the representation of the delineation of visual cortical regions per subject, analyzed and visualized in the *mrVista* program. Here, the anatomical preprocessing, gray and white matter segmentation and the visualization of the visual cortex, were already analyzed and the signature experiment, outlined in the following sections, were imported in this framework.

Revealing the signature integration fields used exactly the same preprocessing

---

<sup>1</sup>Second order-statistics are more discriminating for texture analysis than mere first-order statistics. It seems to be unessential to use statistics of higher order than of order 2, because it has been demonstrated that it is very hard for humans to discriminate images with identical statistics of order one and two but different higher order (Julesz, 1975).

protocols as described in the visual field mapping experiments. Hence, the same materials and methods as that described in the visual field mapping experiment for; subjects, Magnetic Resonance Imaging, functional preprocessing and stimulus presentation were used. However, the alignment of the functional data relies on the transformation of the inplane anatomy, recorded in the signature scan sessions, to the high-resolution anatomy obtained in the visual field mapping experiments. The following sections describe deviant methods like the center-surround localizer, stimulus description, experimental paradigm and the methods to reveal the signature in response to the performed experiments.

### 4.2.2 Stimulus description

The visual stimuli used in the visual field mapping are created using Matlab (MathWorks, Natick, MA) and supplementary tools provided in the Psychtoolbox (Brainard, 1997; Pelli, 1997).

#### *Texture-like stimuli*

The texture-like stimuli are constructed from many orientational bars. The reason for this is that orientational bars are known to elicit neural activity of simple cells and the statistical properties of many orientational bars are easy to manipulate. Furthermore, Parkes et al. (2001) have shown, in an orientation discrimination task, that subjects readily report the average orientation of the target and flankers that fall within the critical region. The orientational bars in the texture-like stimuli are constructed from two-dimensional Gabor filters (Lee, 1996) to sustain a mean luminance for the entire screen. Hence, the measured fMRI-signal cannot be due to luminance changes in the screen and frequency selective processes in early vision are stimulated (Jones et al., 1987). Each Gabor patch is a windowed sinusoidal luminance pattern and in space domain is determined as follows

$$g(x, y) = s(x, y) w_r(x, y)$$

where the Gabor patch  $g(x, y)$  is a complex sinusoid function, i.e. Gabor carrier,  $s(x, y)$  multiplied by a Gaussian envelope  $w_r(x, y)$ . The Gabor carrier  $s(x, y)$  is defined as

$$s(x, y) = L(1 + C \exp(j(2\pi(u_0x + v_0y) + P)))$$

where  $L$  is the background luminance,  $C$  the contrast of the Gabor patch,  $(u_0, v_0)$  is the spatial frequency and  $P$  the phase of the Gabor carrier. The spatial frequency is set to complete two periods in the total size of each Gabor patch. The actual size is

reshaped from a  $100 \times 100$  pixel grid and therefore the frequency is set  $100/2 = 50$  pixels/cycle. In the signature experiments the actual size of each Gabor patch is  $14 \times 14$  pixels. A two-dimensional circular Gaussian envelope is then multiplied with the carrier to smooth the Gabor patch. The envelope is computed with the following equation

$$f(x, y) = A \exp\left(-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}\right)$$

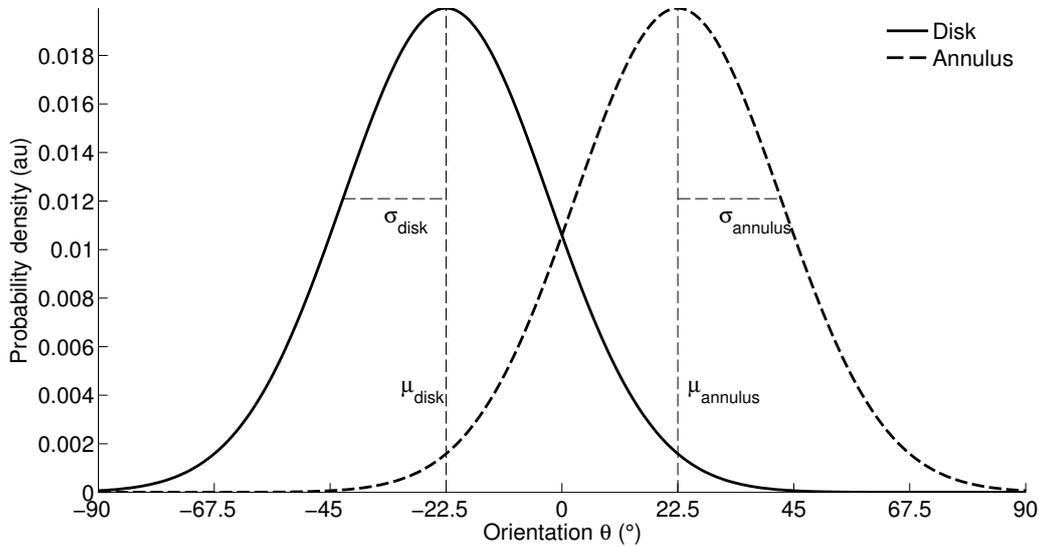
where  $A$  is the amplitude of the envelope and  $x_0$  and  $y_0$  the center in the pixel grid of the two-dimensional Gaussian distribution with respectively a spread of  $\sigma_x$  and  $\sigma_y$ . The size of a Gabor patch is uniform in vertical and horizontal direction. Therefore, the spread of the two-dimensional envelope is identical when defining a circular envelope, i.e.  $\sigma_x = \sigma_y$ . Furthermore, the Gabor patches are 20 pixels spaced from each other and are only presented inside the maximum radius of the screen. Hence, the height of the screen determined the maximum radius wherein Gabor patches are drawn. The reason for this is that the Gabor patches are evenly distributed along eccentricity. In other words, integration fields contain the same information near the borders of the screen.

The illusory contour is defined at the transition between the disk and annulus and is determined at half the maximum radius (height of the screen). In figure 1.4 the orientational properties of Gabor patches in either the disk or annulus are drawn from distinct uniform normal distributions. In figure 4.1 the Gaussian probability density functions of the orientational properties for the disk and annulus are plotted, where  $\Delta\mu = |\mu_{disk} - \mu_{annulus}| = 45^\circ$  and  $\sigma = \sigma_{disk} = \sigma_{annulus} = 20^\circ$ . In other words, the angles of the Gabor patches in the disk and annulus are drawn from a similar normal distribution except for their difference in mean orientation. The illusory contour is prominent when the orientational properties of the oriented Gabor patches in the disk and annulus differ substantially.

To allow a signature of neural integration processes the difference between the mean orientation of Gabor patches contained in the disk or annulus was modulated. In the following section the modulation of the statistical properties between the disk and annulus is outlined.

#### *Stimulus modulation*

In the signature experiments, the difference between the mean orientation of the central disk and surrounding annulus is modulated with a specified frequency. The maximal difference between the mean orientation of the disk and annulus was set



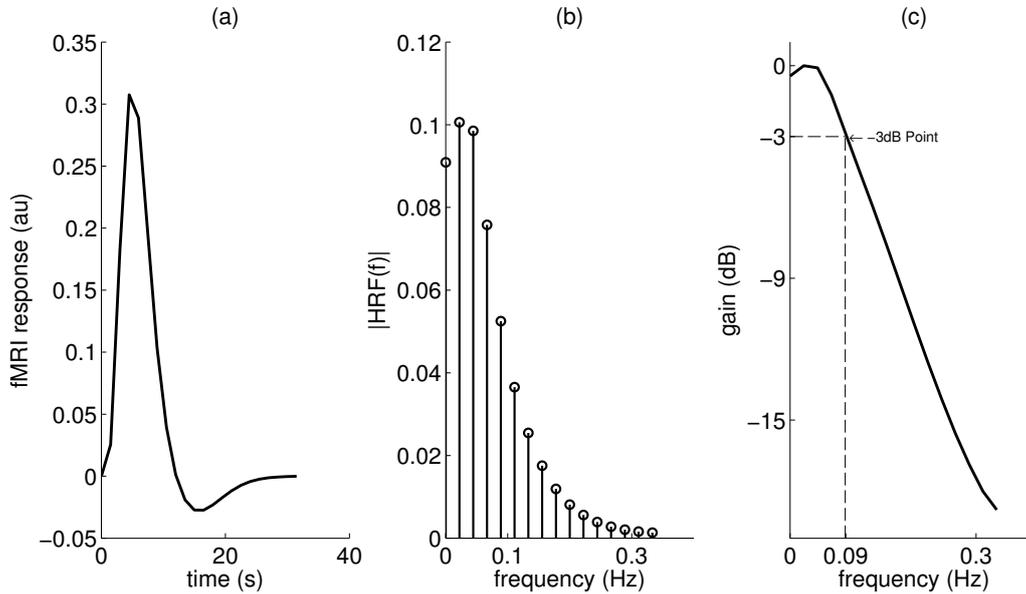
**Figure 4.1:** *Distribution of Orientational Stimuli. Normal probability density functions of the orientational properties of Gabor patches contained in the center disk or surrounding annulus disk. Only the mean orientation of the distributions differ.*

to  $\Delta\mu = 45^\circ$ . Furthermore, the standard deviation of the orientational properties of the Gabor patches in the disk and annulus are the same, i.e.  $\sigma = \sigma_{disk} = \sigma_{annulus}$ , where  $\sigma = 20^\circ$  for the second-order condition and  $\sigma = 0^\circ$  for the lateral inhibition condition.

The position and orientations of the bars that form the pattern randomly rotate and jump continuously in a rate of  $8 \text{ Hz}^2$  to account for adaptation (Boynton & Finney, 2003), oblique effects (Furmanski & Engel, 2000) and center-surround interactions (Xing & Heeger, 2000). Thus, for each frame, the position of the Gabor patches randomly alter maximal 4 pixels in horizontal and vertical position from their origin. Furthermore, for each frame a random mean orientation is chosen for the Gabor patches inside the disk. Hence, the mean orientation of Gabor patches in the annulus differ the amount that is determined by the modulation.

Constraints to the experimental paradigm used for the modulation exist. For instance, the hemodynamic response function (HRF) acts like a low-pass filter (for more information on the HRF see Appendix A.1.2). Therefore, to sustain enough power of the modulation frequency in the fMRI-signal, the frequency should generally be lower than the half-power point of the HRF filter properties., i.e. lower

<sup>2</sup>This is slow enough to allow for figure/ground separation for each frame. Moreover, it cannot be due to temporal feature integration because the (second order) figure/ground differences are defined in the spatial domain, not in the temporal domain.



**Figure 4.2:** *HRF Low-Pass Filter Properties.* (a) The hemodynamic response function (HRF) as function of time, (b) in the frequency domain (single-sided amplitude spectrum) the HRF looks similar to a low-pass filter. Consequently, (c) frequencies beyond the half-power point of the HRF are substantially filtered out.

than a  $-3dB$  power gain. This point can be found with the following equation

$$Gain(dB) = 10 \cdot \log(I_1/I_0)$$

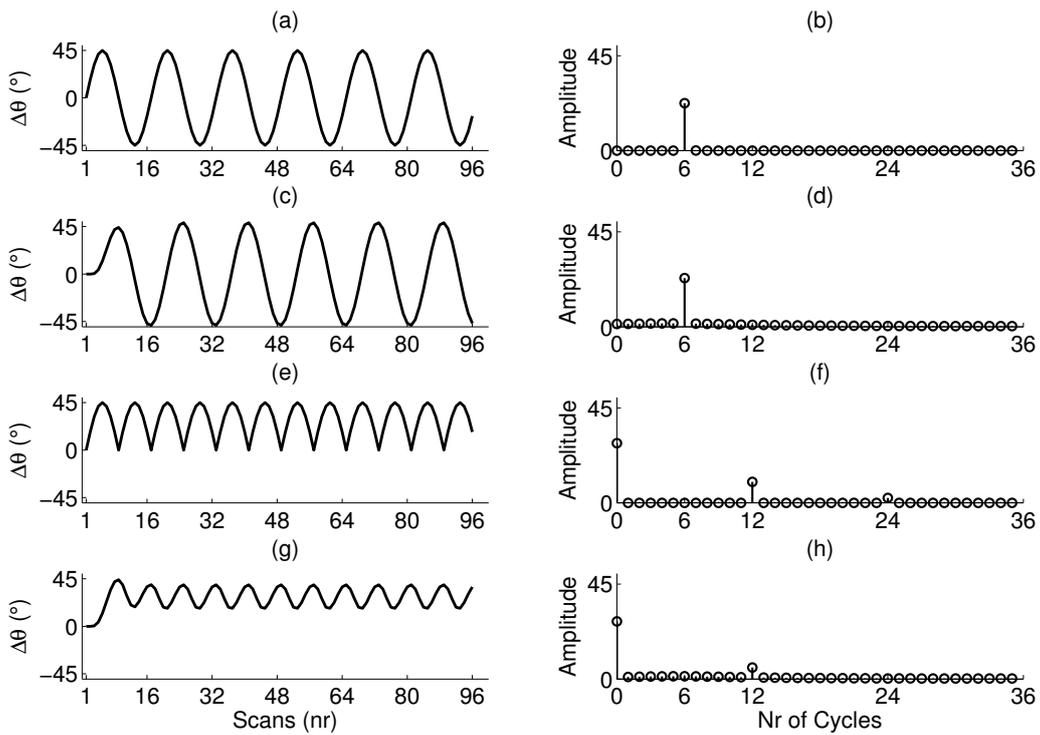
where  $I_1$  is the amplitude spectrum of the HRF and  $I_0 = \max(I_0)$ , i.e. the maximum amplitude of the HRF spectrum. See figure 4.2 for the HRF filter properties and the  $-3dB$  power gain. The frequency of the modulation should be roughly below  $0.09Hz$  corresponding with 12 cycles per run in a standard experiment paradigm of 144 seconds. To be sure to get the highest activation a modulation of 6 cycles per run is used. Hence, the time of a single cycle is 24 seconds corresponding with 16 TRs.

The modulation of the difference between the mean orientation of Gabor patches in the disk or annulus was first done with a sinusoid with an amplitude of  $45^\circ$ , see figure 4.3a. The frequency of the modulation, as explained above, is determined at 6 cycles per run. A problem arises because the first half period is perceptually the same as the second (subjects notice the illusory contour but not register a positive or negative difference between the mean orientation). Therefore, a rectified sinus could be part of the time series measured in the fMRI data (figure 4.3e). In the

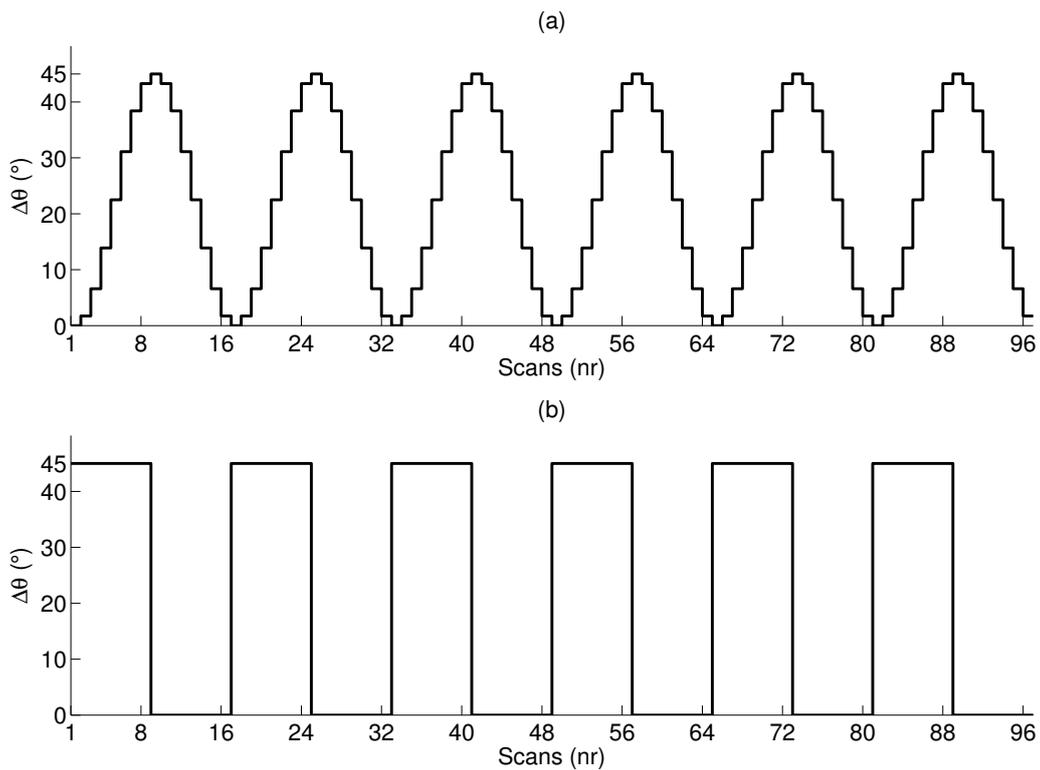
amplitude spectrum of the rectified sinusoid a peak appears at two times the modulation frequency, i.e. 12 cycles per run, and additional harmonics. Consequently, the filter properties of the HRF has a substantial influence on the rectified sinusoid in contrast to the input sinusoid.

Now, a positive sinusoid between zero and the maximum orientation difference can be used with a frequency of 6 cycles per run, see figure 4.4a. This results in a very clear amplitude spectrum. However, the duration of the perceptually presence illusory contour is uncertain since the underlying structures could process only the presence of an illusory contour and the time series of the fMRI-signal is a block paradigm. Then, the temporal extent of this block perception is uncertain.

To be certain about the perceptual experience, a block paradigm can be used with a frequency of 6 cycles per run, see figure 4.4b. In this paradigm the first half period the mean orientation is maximal and in the remaining half period there is no mean orientation difference between the disk and annulus. Therefore, in the first half period an illusory contour is perceived and the second half period there is no illusory contour. Both, the positive sinusoid and the block paradigm, are recorded in the signature experiments. See table 4.2 for the exact recordings of the signature lateral inhibition (SLI) condition ( $\sigma = 0^\circ$ ) and signature mean (SM) condition ( $\sigma = 20^\circ$ ).



**Figure 4.3:** *Time Series of Rectified Sinus Modulation.* (a) The initial modulation of the mean orientation of the Gabor patches contained in either the disk or annulus. The modulation is a sinusoid with a frequency of 6 cycles per run and an amplitude of 45 degrees. (b) The single-sided amplitude spectrum of the initial experiment. The input sinusoid modulation convolved with the HRF (c) results in a similar profound single-sided amplitude spectrum. (e) If the perceptual experience is taken into account, a rectified sinusoid represents the properties of the modulation in the time series of the fMRI-signal. (f) The single-sided amplitude spectrum of the rectified sinusoid shows a peak at double the frequency of the input sinusoid and consists of harmonics. (g) When the rectified sinusoid is convolved with the HRF, (h) the expected time series in the fMRI signal will show a lower peak in the amplitude spectrum due to the low-pass filter properties of the HRF. Furthermore, the harmonics are completely filtered out.



**Figure 4.4:** *Time-Series of Modulation. (a) The positive sinusoid modulation of the mean orientation. Every TR the difference of the mean orientation advances according to a sinusoid with a frequency of 6 cycles per run. (b) Another used experimental paradigm consist of a block signal with a frequency of 6 cycles per run. Here, the mean orientation difference is maximal in the first-half period and zero in the second-half of the period.*

Experiment	stimulus period (sec)	NR of cycles	NR of runs	Total duration (min)
Session 1: SLI Block	24	6	4	9.6
Session 2: SLI Sinusoid	24	6	4	9.6
Session 3: SM Block	24	6	4	9.6
Session 4: SM Sinusoid	24	6	4	9.6
Session 5: CS-Localizer	24	6	3	7.2

**Table 4.2:** *Signature Integration Fields Experiments.*

### 4.2.3 Center-Surround Localizer

To map the illusory contour onto the visual cortex, a center-surround localizer was used. The stimuli are presented on a mean luminance background and consisted of contrast moving checkerboard patterns (black and white), radially drifting from or away from fixation point with 8 frames per second. The checkerboards are in shape of the disk (center) and annulus (surround). These two shapes are presented to the subject sequentially for a duration of 8 TRs each. Thus a full cycle is 16 TRs. A cycle is repeated 6 times in a run and therefore total scan duration of a single run is  $16 \times 6 \times 1.5 = 144$  seconds. The center-surround localizer is performed 3 times to get a good signal to noise ratio (see table 4.2).

Then, a traveling-wave analysis was performed to delineate the cortical regions representing the disk and annulus. Hence, a Fourier analysis of the time series in every voxel in the visual cortex reveals the phase in response to the fundamental frequency of the stimulus modulation. The transition between the phase is the location of the transition in the visual cortex. The transition was manually determined in a flattened view of the visual cortex.

### 4.2.4 Signature mapping

Regions of interest (ROIs) were determined for the center and surround parts for the primary (V1) and extrastriate cortical region (V2) in the left and right hemisphere for each subject. Furthermore, a ROI is determined for the illusory contour. The ROIs were determined in a flattened view of the visual cortex. Then, distances to the illusory contour of every voxel in V1 and V2 were calculated in three-dimensional cortical space along the boundaries of the white matter segmentation. The illusory contour is at distance zero and the distances of voxels towards the fovea are negative whereas voxels towards the periphery are positive. To summarize, the cortical distance to the illusory contour of every voxel in the visual cortex were analyzed and scaled negative or positive depending on their location (disk or annulus).

Now, voxels are binned according to the cortical distance to the illusory contour with an equally spaced distance of 2.5 millimeter (center of bin is determined at 0 millimeter and advances 2.5 millimeter for each bin in positive and negative distance). In other words, the time series of voxels that fall within the same cortical distance bin are averaged. The binned voxel responses allowed a signature plot, i.e. responses to the modulation of the mean orientation (integration processes) as function of cortical distances to the mapped illusory contour. Here, the analysis is restricted to V1 and V2. The visual cortical regions are averaged, i.e.

left and right hemisphere and additionally for V2 the upper and lower quadrant are averaged.

The standard error of the mean for every bin, was calculated with the pooled standard deviation. This a better estimate of the average of the single standard deviations. Assuming independent measurements<sup>3</sup> the pooled standard deviation of  $k$  measurements is calculated as follows

$$SD_{pooled} = \sqrt{\frac{(N_1 - 1)SD_1^2 + (N_2 - 1)SD_2^2 + \dots + (N_k - 1)SD_k^2}{N_1 + N_2 + \dots + N_k}}$$

where  $N_i$  is the number of samples, average is  $\bar{x}_i$  and standard deviation is  $SD_i$  of measurement  $i$ . Then, the uncertainty in the averages, i.e. standard error of the mean ( $SEM$ ), is

$$SEM = SD_{pooled} \sqrt{\frac{1}{N_1} + \frac{1}{N_2} + \dots + \frac{1}{N_k}}$$

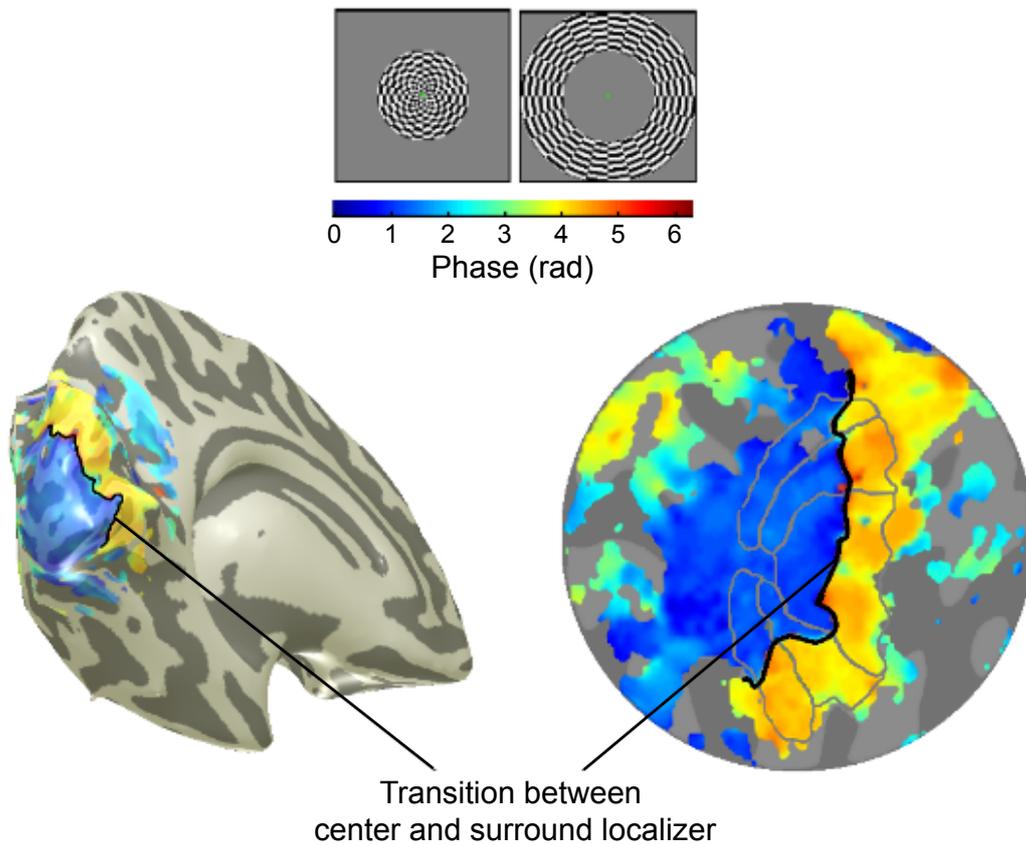
### 4.3 Results

The visual cortex that represents the disk, annulus and the transition between the disk and annulus was determined with the functional responses to the center-surround localizer. The phase of the traveling-wave analysis of the time series of every voxel in response to the localizer reveals the spatial correspondence between the stimulus and location of a voxel. For subject KH, the results of the center-surround localizer are shown in figure 4.5. The phase is color coded in the legend of the figure. A clear distinction between voxels that represent either the disk or annulus is prominent.

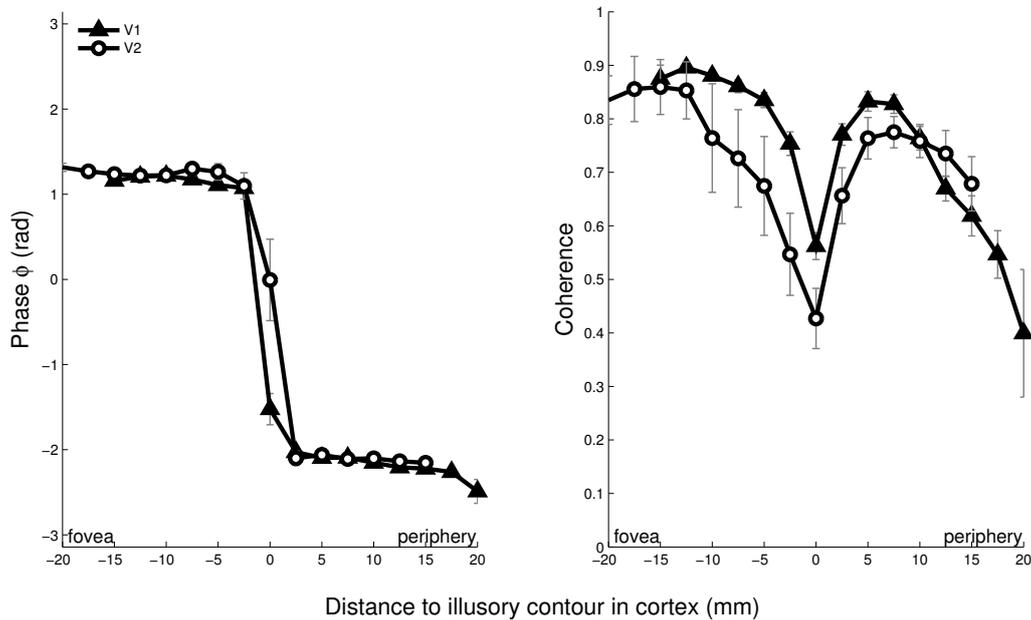
The analysis of the center and surround in every cortical region allows the computation of the cortical distances to the illusory contour. Regions of interest (ROI) were determined for the transition between the center and surround and the mapped center and surround representation in cortical regions V1 and V2. The distinction between center and surround was necessary to account for distances towards the fovea (disk) or periphery (annulus). Cortical distances of voxels representing the center or surround were respectively scaled negative and positive. Furthermore, the voxels were binned with a bin size of 2.5 millimeter that allows for plotting a signature.

---

<sup>3</sup>Neighboring voxels are likely to depend on each other. Therefore, the standard error of the mean is an approximation.



**Figure 4.5:** *Results Center-Surround Localizer. The upper part of the figure shows the phase that corresponds with the center or surround stimulus presentation. The phase of every voxel in the visual cortex in response to the localizer experiment is projected onto an inflated mesh and flattened cortex. The transition between the phase of the center and surround is the representation of the illusory contour in the visual cortex.*



**Figure 4.6:** *Signatures Center-Surround Localizer.* In the left part of the figure, the binned phase responses are plotted with respect to cortical distances to the transition. There is a clear transition of phase responses at the transition (distance is 0 mm). The coherence signature is plotted in the right figure. Here, there is drop of coherence at the illusory contour. This is obvious considering that the voxels near the illusory contour respond to both the center and surround checkerboard stimulus. The error-bars show the standard error of the mean.

For instance, the responses to the center-surround localizer can be plotted as a signature. In figure 4.6 the phase and coherence signature of the center-surround localizer experiment is plotted for V1 and V2. Distances beyond 15 millimeter are not plotted, because of the absence of strong responses (especially in the signature using texture-like stimuli). However, we are only interested in responses that are close to the transition. The left part of the figure, i.e. signature of the phase, shows a clear transition of the phase of the center and surround checkerboard stimulus presentation. The right part of the figure, i.e. coherence signature, shows a drop in coherence of voxels that are close to the transition. This can be explained by the fact that the population receptive fields of these voxels overlap both the center and surround stimulus. Hence, these voxels respond to both the center and surround checkerboards.

There are two main factors that influence the phase signature of the center-surround localizer; the delayed peak response of the BOLD-response in a active

voxel time course, i.e. the temporal properties of the HRF, and the underlying phase of the stimulus modulation. A simplified analysis of the phase is that the peak response of the HRF is approximately at 6 seconds and this corresponds with a phase delay of  $6/24 = 1/4$  of a cycle, i.e. a phase delay of  $2\pi \cdot 1/4 = \pi/2$  radians. Furthermore, the center and surround checkerboard stimuli are respectively presented in the first- and second-half of a cycle corresponding with a radial phase delay of 0 of the center and  $\pi$  of the surround induced BOLD modulation. Hence, the total phase delay of the center stimulus is  $\pi/2 + 0 = \pi/2$  and of the surround stimulus is  $\pi/2 + \pi = 3\pi/2$ . In the phase signature plot of the center-surround localizer experiment shows roughly similar phases of voxels representing the center and surround stimulus presentation.

The signatures of integration processes, i.e. responses to the modulation of the texture-like stimuli, were analyzed and plotted according the same strategy. Thus, the signature of integration processes are analyzed with the traveling wave method which provides phase, coherence and amplitude information. We are especially interested in the amplitude of neural integration processes. The amplitude signature reveals the amplitude of the peak response of the modulation frequency as function of cortical distance to the illusory contour. In figure 4.7(a) the amplitude signature is plotted for each condition and modulation paradigm, i.e. the block and sinusoid modulation of the signature lateral inhibition (SLI) condition and signature mean (SM) condition. There is no clear signature visible that resembles neural integration process correlates.

This is not discouraging because the responses depend on the phase as well since the largest response of a voxel can be out of phase compared with the stimulus modulation phase. In figure 4.7(b) the signature of the phase is plotted. Here, the phase of the responses are averaged using a complex notation which overcomes the problem of averaging a phase of 0 and  $2\pi$  to  $(0 + 2\pi)/2 = \pi$ . The signature of the phase shows that voxels near the illusory contour are responsive in a constant phase in contrast with voxels further away from the illusory contour. To account for these phase shifts of the responses, we used the so called projected amplitude. The projected amplitude assures a fixed temporal correspondence between the stimulus modulation and the BOLD-response modulation. Hence, the projected amplitude corrects the amplitude of the BOLD-response for every voxel to a fixed phase.

The projected amplitude for every voxel is calculated as follows

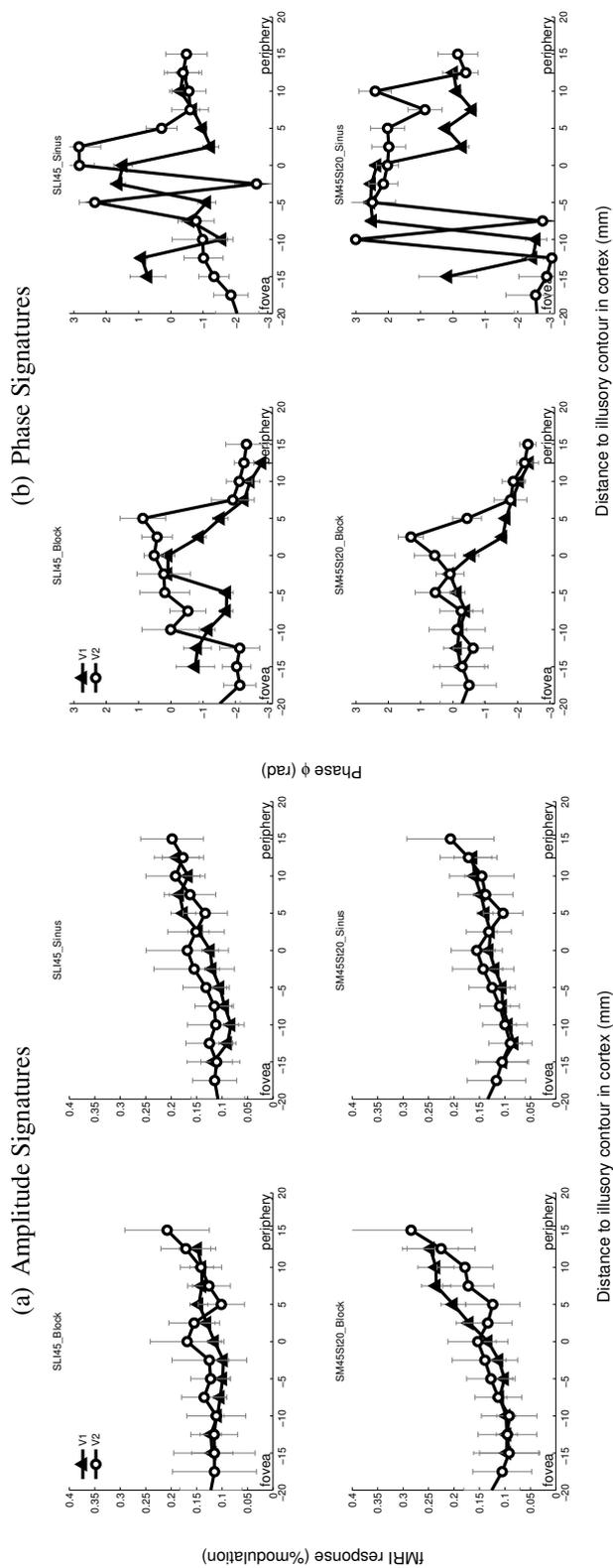
$$A_{proj} = A \cos(P - P_{sync})$$

where  $P$  is the phase of the BOLD modulation at the fundamental frequency of

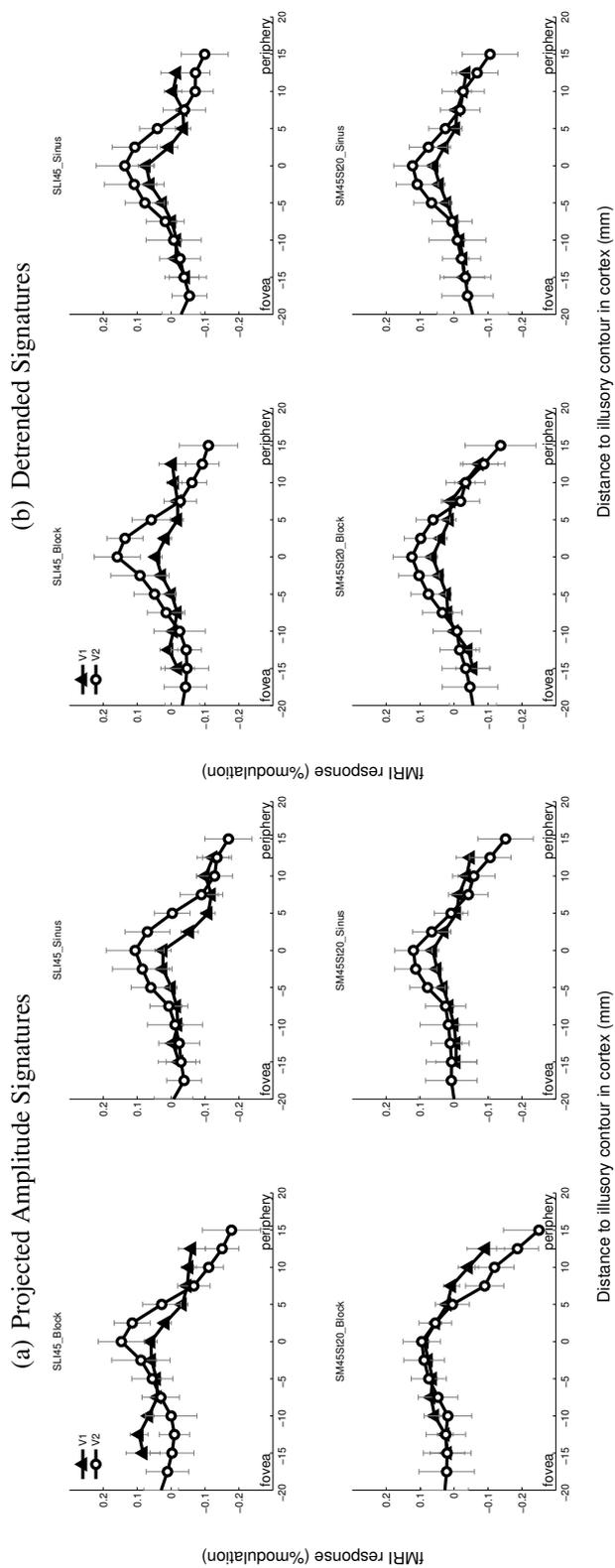
the stimulus and  $P_{sync}$  is the synchronization phase. The synchronization phase assures a fixed temporal correspondence between the stimulus and BOLD modulation. For example, when a BOLD modulation is in correspondence with the underlying stimulus modulation,  $A_{proj} = A$  because  $P = P_{sync}$ . The synchronization phase  $P_{sync}$  is determined as the phase at the illusory contour since integration fields near the illusory contour are the most responsive to the stimulus modulation (first a theoretical approach was used that was not satisfactory, see Appendix C.1). Therefore, the phase of the BOLD modulation at the illusory contour is used as the synchronization phase. In figure 4.8(a) the resulting signature is plotted. Here, a bell shaped curve with its center around the illusory contour is prominent.

A trend with a negative slope in all conditions is visible. From our stimulus design we would not expect this effect. Therefore, the precise processes underlying this effect remains unclear but could be attributed to vascular (blood) stealing (Woolsey et al., 1996) or the influence of the negative BOLD-signal (Harel et al., 2002; Shmuel et al., 2006). Blood stealing is the influence of large blood vessels on the fMRI signal. The neural link to negative BOLD signals remains unclear. Other considerations for the possibility of the trend is the influence of anisotropy or the maximum stimulus presentation. Anisotropy resembles a stronger crowding effect towards the periphery and could cause a decrease in the BOLD-signal. The maximum stimulus presentation prohibits a stimulus presentation outside its borders. Therefore, integration fields towards the stimulus presentation borders contains a decreasing number of Gabor patches and could be a possibility for the trend. Both possibilities were refuted because these cannot be attributed to the stimulus modulation. In other words, the stimulus presentation in each condition always consisted of the same amount of Gabor patches. So, no differences in the signature can be attributed to the presentation of the single Gabor patches or other influences that are independent to the stimulus modulation.

To be able to get a grip of the found signatures a linear detrend of the projected amplitude is performed by removing the least squares fit of a straight line through the data points (least-squares fit). The results are plotted in figure 4.8(b). When looking at the signature plots of the sinusoid and block modulation paradigm, we can conclude that there are no significant differences. This is obvious considering that the modulation paradigm does not facilitate distinct neural processes. Furthermore, there are no clear differences between the SLI and SM condition. The presented data are averaged data of two subjects. For individual data see Appendix C figure C.2 and C.3.



**Figure 4.7:** Signatures of Amplitude and Phase. The signatures were analyzed in V1 (diamonds) and V2 (circles) for every condition, where in (a) the amplitude of BOLD-modulation is plotted as function of cortical distance to the illusory contour and in (b) the phase of the BOLD-modulation is plotted as function of cortical distance to the illusory contour.



**Figure 4.8:** Signatures of Projected Amplitude and Detrended Projected Amplitude in V1 and V2 voxels and are plotted for every condition. In (a) the projected amplitude and in (b) the detrended projected amplitude is plotted as function of cortical distance to the illusory contour.

## Signature Integration Fields: a Theoretical Approach

Finding the neural correlates of integration fields in the visual cortex is the prominent research goal of this thesis. Chapter 4 outlined the search of these neural correlates by extracting a neural signature of orientational integration processes using fMRI. The texture-like stimuli used in the experimental paradigm define an illusory contour between a disk and annulus that is caused by distinct statistical summaries caused by orientational properties differences of oriented Gabor patches. Such texture-like stimuli are especially used because cells in the early visual cortex are mainly responsive to edges and contours.

A neural biological plausible model is proposed by van den Berg et al. (2010) that explains crowding as hardwired feature integration using oriented Gabor patches. The proposed model reproduces a wide variety of psychophysical results such as the size of the critical region, compulsory averaging and other discrepancies of the crowding phenomenon. Therefore, supporting the prevailing notion that the underlying mechanism of crowding is 'hardwired' feature integration. With this model, a virtual fMRI experiment was performed that allowed the comparison between the signature of integration processes found in fMRI and the crowding model. This chapter is an outline of the crowding model and the performed virtual fMRI experiment.

### **5.1 The Crowding Model**

In Oliva et al. (2004) visual complexity of a scene is represented as multiple stimuli properties such as symmetry, organization and clutter. Scenes with high den-

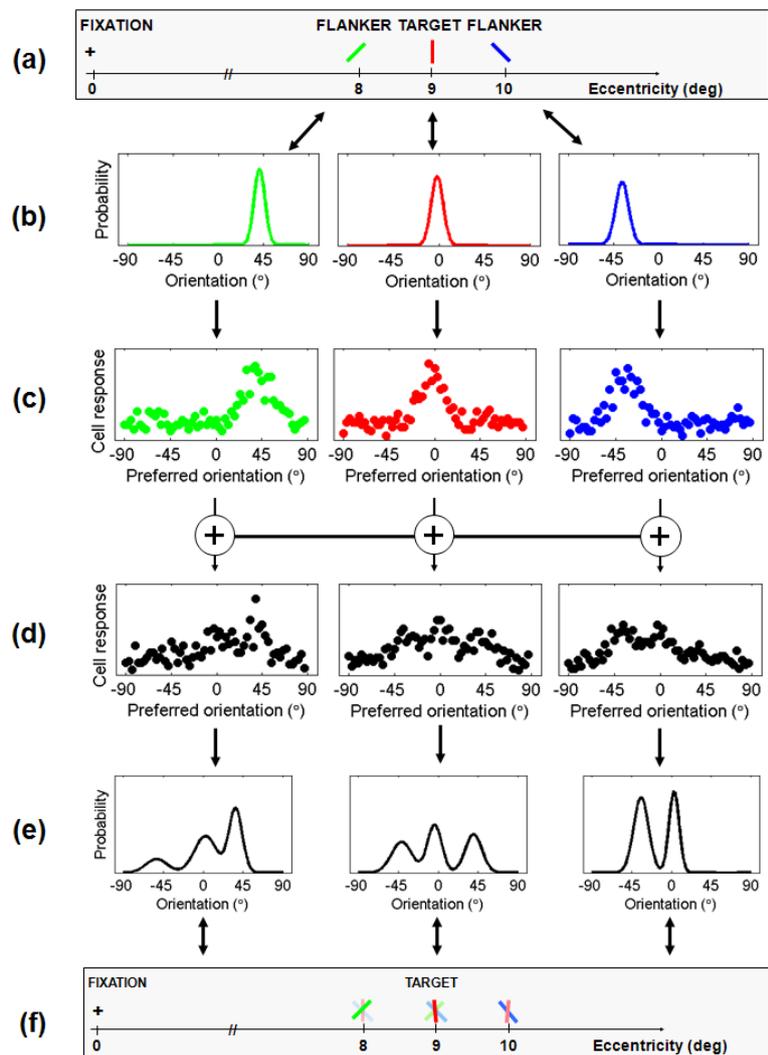
sities or chaotic visual information are cluttered. Crowding is closely related to clutter (van den Berg et al., 2009) and poses a neural information processing problem which is well explained by the principles of population coding (Pouget et al., 2000). Population coding provides a mathematical and biological plausible solution for the encoding of visual stimuli by population of cells.

In van den Berg et al. (2010) a population coding model, based on the distributional population coding (DPC) scheme (Zemel et al., 1998) is proposed to model the underlying mechanism of the crowding phenomenon. For example, the model respond to crowded stimuli as depicted in figure 5.2. The model consist of two layers. The first layer encodes single bars presented at each location in the visual field. Hence, orientation selective neuronal populations encode the orientation of single bars and results in a probability distribution of orientation. In the second layer these orientational probability distributions are pooled with a weighted two-dimensional Gaussian overlay distribution to model feature integration. The spatial extent of the Gaussian overlay function resembles the size of integration fields. This weighted summation is referred to compulsory averaging and is a general property found in crowded orientation discrimination tasks. For instance, subjects reported the average orientation of crowded orientational Gabor patches where the orientation discrimination of individual stimuli is impaired (Parkes et al., 2001). In the following sections a simplified mathematical outline of the first and second layer of the model is presented.

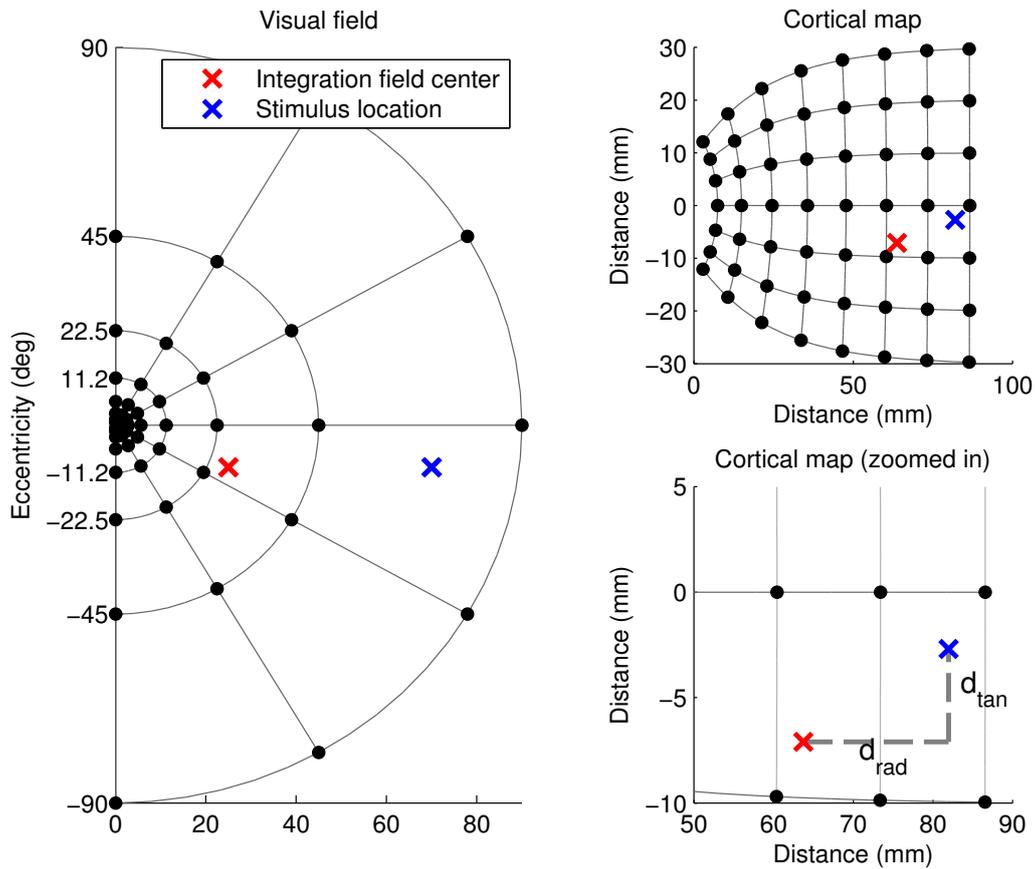
### 5.1.1 Stimulus encoding

In this section a simplified mathematical outline of the first layer (stimulus encoding) of the crowding model is presented. Input stimuli are defined as  $S = (\theta, \alpha, \lambda, c)$  parameter tuples, where  $\theta$  is the orientation,  $\alpha$  size of the Gabor patches,  $\vec{\lambda} = (\lambda_x, \lambda_y)$  the location in the visual field and  $c$  the relative contrast of the stimulus. Each input stimuli is mapped onto visual cortical space using a complex log mapping introduced in Schira et al. (2007). See figure 5.2 for a graphical illustration of the mapping of visual field to cortical space. Before encoding the orientation of the bars with population coding principles, stimulus uncertainty caused by neural noise is modeled as an probability distribution of the stimulus orientation.

Now, encoding a stimulus at a single location follows the extended Poisson model. Neuronal activity is noisy and therefore it is convenient to use a probabilistic view on the encoding paradigm. Hence, the model assumes that information of the stimulus is noisy and not fixed. The average neural activity of a cell is



**Figure 5.1:** *Biological Plausible Model of Crowding.* The model consists of several steps. (a) The visual field consist of three orientational bars. (b) The encoding of angle information of the individual stimuli are subjected to neural noise which causes uncertainty. (c) Then, the first layer of the model encodes the probability distribution of the individual stimuli into neural spike rates according to population coding techniques. (d) Then, the second layer pools the encoded probability distribution, where integration fields are centered onto each bar stimulus, . This result in three probability distributions of the pooled orientation of the bars. (e) A mixture of normal distributions are used to decode the pooled probability distributions. (f) The percept at each location is altered due to the crowding effect. Hence, there is no clear distinction of the targets orientation, i.e. the target is crowded with the surrounding flankers and results in a jumbled representation of the oriented bars at a certain location. (source: van den Berg et al. (2010))



**Figure 5.2:** Illustration of Mapping Visual Field to Cortical Space Locations. (a) Right visual field location marked as black filled circles are transformed to cortical space locations shown in (b). This transformation is computed using a complex log mapping introduced in Schira et al. (2007). The distance of a stimulus, marked as a blue cross, to an integration field center, marked with a red cross, is computed in cortical space. (c) A zoomed illustration of how the distance is calculated. The radial distance is the distance along the eccentricity axis and the tangential distance is the distance along the orthogonal axis (source: van den Berg et al. (2010))

numerically computed as follows

$$\langle r_i \rangle = r_{base} + \sum_j \Theta_j F_{ij}$$

where  $r_{base}$  is spontaneous activity of a cell,  $\vec{\Theta} = \{\Theta_1, \dots, \Theta_j\}$  the input distribution (noisy angle representation) approximated by histograms and  $\vec{F} = \{F_{1j}, \dots, F_{iJ}\}$  are the tuning curves approximated by histograms for each cell (tuning curves model orientation selective cells). Both histograms have equally spaced bins in the range  $[-\pi, \pi]$ . Then, the population code  $\vec{r}_h = \{r_{h1}, \dots, r_{hJ}\}$  of a stimulus tuple  $S_h = (\Theta_h, \alpha_h, \vec{\lambda}_h, c_h)$  is constructed from Poisson distributions

$$P[r_{hi} | \vec{\Theta}_h] = e^{-\langle r_{hi} \rangle} \frac{\langle r_{hi} \rangle^{r_{hi}}}{r_{hi}!}$$

### 5.1.2 Stimulus integration

Stimulus integration is computed in the second layer of the crowding model. An integration field is modeled as a two-dimensional weight overlay function. When  $\vec{\lambda}_h$  represents the position of stimulus  $h$  and  $\vec{\lambda}_k$  the centre position of integration field  $k$ , the weight overlay function is determined as

$$w(\vec{\lambda}_h, \vec{\lambda}_k) = \exp\left(-\frac{[d_{rad}(\vec{\lambda}_h, \vec{\lambda}_k)]^2}{2\sigma_{rad}^2} - \frac{[d_{tan}(\vec{\lambda}_h, \vec{\lambda}_k)]^2}{2\sigma_{tan}^2}\right)$$

where  $\sigma_{rad}$  and  $\sigma_{tan}$  is the size in radial and tangential direction of an integration field, and  $d_{rad}$  and  $d_{tan}$  are respectively the radial and tangential distance to the cells in cortical space (example in figure 5.2). This determines the strength of crowding as function of cortical distance. In other words, the cortical distance between the center of the integration field and the center of each stimulus determines the strength of crowding. The integration of population codes in an integration field is then captured in the equation

$$R_{hi} = \sum_{k=1}^N w(\vec{\lambda}_h, \vec{\lambda}_k) r_{ki}$$

where  $N$  is total number of input stimuli  $w(\vec{\lambda}_h, \vec{\lambda}_k)$  resembles the weight overlay function in cortical space and  $r_{ki}$  are the encoded population codes determined in

layer one. To summarize, orientation integration is modeled as a two-dimensional Gaussian weight overlay function, resembling the processes in an integration field, multiplied with the activity of the individually encoded stimuli. The resulting summation is a population coding distribution of the net-response of an integration field.

## 5.2 Materials and Methods

In this section the methods to investigate the signature of integration processes using the biological plausible model for crowding is explained. The aim of this study is to compare the results of this theoretical approach with the performed fMRI study outlined in chapter 4. Therefore, exactly the same methods and restrictions of the fMRI study were applied.

### 5.2.1 Stimulus description

The texture-like stimulus is constructed in the same manner as that described in section 4.2.2. However, every Gabor patch was represented in the parameter tuple  $S = (\theta, \alpha, \lambda, c)$ , where size and location are represented in visual field space. Hence, the amount of pixels  $P$  and visual angle  $A$  of the stimulus presentation correspond with  $P_{width} : P_{height} = A_{width} : A_{height}$  that is similar to  $800 : 600 = 32.7^\circ : 24.8^\circ$  (see section 3.2.7).

The virtual fMRI experiment should allow the comparison between the crowding model and the fMRI experiment. Therefore, the texture-like stimuli in the experimental paradigm of the fMRI experiment resemble net responses of integration fields. Thus, a signature plot of the virtual fMRI experiment should follow the same techniques. Each frame in the texture-like stimuli is modeled according to the crowding model. Hence, every Gabor patch presented on the screen is encoded in the first layer of the model. Then, the location of integrations fields are determined with respect to the illusory contour. The location of the illusory contour, defined at half the maximum eccentricity radius of the screen, is mapped onto cortical space. From here the center of the modeled integration fields are 2.5 millimeter spaced from the illusory contour on a horizontal radial line in cortical space with respectively a minimum and maximum of  $-15$  and  $15$  millimeters, i.e.  $d_x = [-15, -12.5, \dots, 0, \dots, 12.5, 15]$  and  $d_y = 0$ .

Only the right visual field is taken into account to reduce computational issues which is possible because no discrepancies exist between integration processes in the left and right visual cortex. Furthermore, only V1 was analyzed and the

cortical magnification of V1 proposed by Schwartz (1980) was used to map visual field locations onto cortical space. Other parameters for the crowding model were set according to the modeled parameters in van den Berg et al. (2010). Thus,  $r_{base} = 5$  spikes/second, the tuning curve width  $\sigma_t$  was  $15^\circ$ ,  $\sigma_{rad}$  and  $\sigma_{tan}$  (size of the integration fields in radial and tangential direction) were respectively set to 2.5 and 1 millimeter and number of neuronal populations  $J$  was set to 90.

### 5.2.2 Signature mapping

For every frame, the stimulus description were encoded using the crowding model and the second layer of the crowding model computed the net-responses of the population coding distributions of the integration fields. Thus, the total amount of stimulus descriptions that were modeled into net-responses of integration fields is  $96 \times 1.5 \times 8 = 1152$ . Furthermore, every condition, first-order and second order induced illusory contour, and block and sinusoid modulation paradigm were investigated. The hypothesis is that the maximum amplitude of these population coding distributions is a measure for the profoundness of the integrated signal. Therefore, for each frame the maximum amplitude of the net-responses were determined for each integration field and frame. The resulting time series of the amplitude is convolved with a standard two-gamma model of the HRF to get the expected BOLD-responses modulation. Then, for each experiment the amplitude of the fundamental frequency in Fourier space reveals the strength of integration processes for each center of an integration field. Hence, revealing the signature of integration processes using a model for crowding.

Furthermore, the time series of voxels are not independent of neighboring voxels due to neural connections and vascular influences (Engel et al., 1997). To account for this artifact a so called linespread function is used to spatially smooth the time series of the modeled integration fields. This linespread function is determined at a full width at half maximum (FWHM) of 3.5 millimeter in the visual cortex and was convolved with the amplitude time series of the virtual fMRI experiment. The FWHM is approximated by a normal distribution

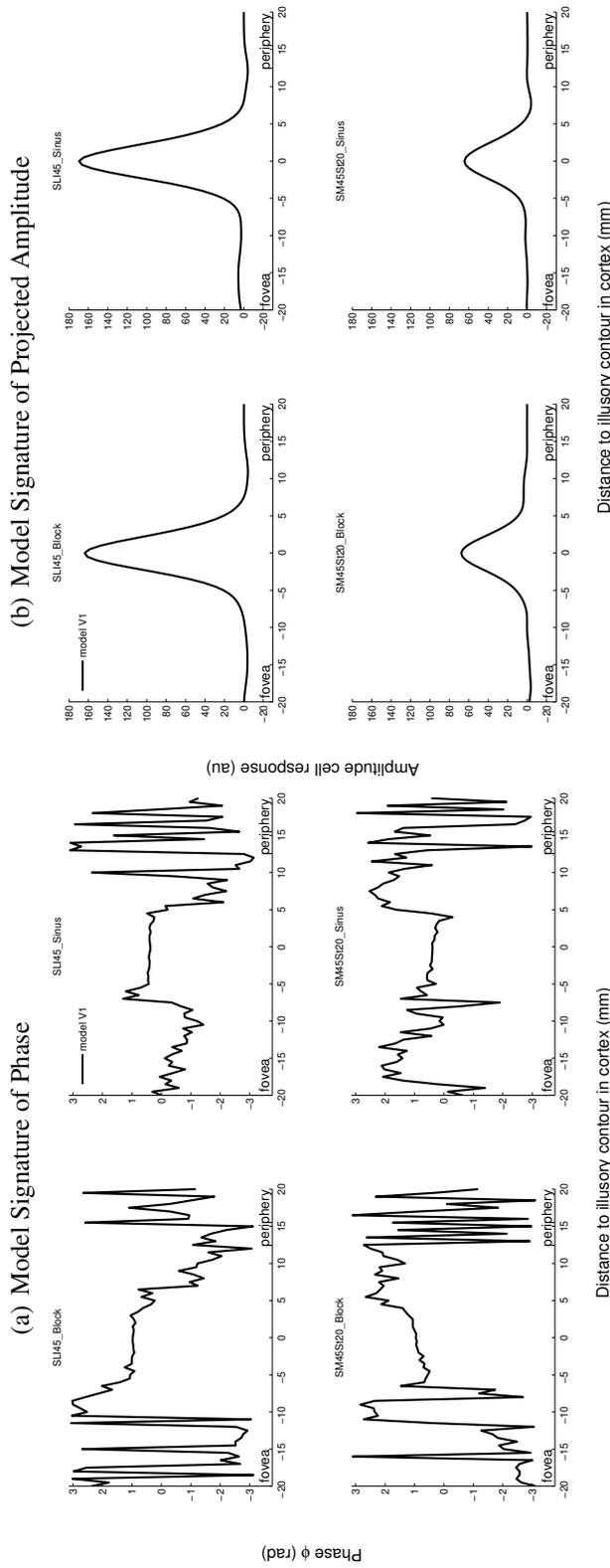
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-x_0)^2}{2\sigma^2}\right]$$

where  $\sigma$  relates to the FWHM as  $\sigma = FWHM/2\sqrt{2\ln 2}$ ,  $x$  is the distance in cortical space and  $x_0$  is its center.

### 5.3 Results

For every experiment, first- and second-order induced illusory contour and the block and sinusoid modulation paradigm, a signature was analyzed for V1. See figure 5.3(a) and 5.3(b) for respectively the signature plot of the phase and projected amplitude for every experiment. The synchronization phase of the projected amplitude is the phase responses at 0 millimeter distance of the illusory contour for every experiment.

The signatures shows a positive bell-shaped curve with its peak centered at 0 millimeter distance to the mapped illusory contour. Furthermore, the amplitude of the SLI condition appears much larger in comparison with the SM condition. This could be explained by bigger fluctuations of the maximum amplitude of the net-responses near the illusory contour in the SLI condition since the stimuli consist of only two different angles. In other words, only two angles are modeled and therefore the net-response of an integration field near the illusory contour shows two high peaks in the amplitude if the illusory contour is visible.



**Figure 5.3:** Signature of Phase. The phase is plotted for every integration field center in V1. The illusory contour is at 0 millimeter and positive and negative distances are receptively towards periphery and foveal representations of the visual field. Only phases near the illusory contour are reliable because no differences in integration fields encapsulating only texture-stimuli contained in the disk or annulus exist in all frames. Signature of Projected Amplitude. The projected amplitude is plotted for every integration field center in V1. The illusory contour is at 0 millimeter and positive and negative distances are receptively towards periphery and foveal representations of the visual field. The signature shows a bell-shaped curve centered on the mapped illusory contour.

## Discussion

A texture, like spatial frequency, is a image region of repetitive structure. For instance, the texture of a table is constant across the table. Such repetitive texture is a visual cue to distinguish the table from surrounding objects. Hence, textures are important in comprehending the gross layout of a visual scene. Image statistics are an important cue for early image segmentation and enable fast processing of texture properties. For instance, early identification of a scene is explained by texture identification models using image statistics (Renninger & Malik, 2004; Bhatt et al., 2007).

It seems that peripheral vision plays an important role in fast processing using image statistics. One of the reasons for this is excessive 'hardwired' pooling of visual features in peripheral vision. This excessive pooling underlies the crowding effect. Although 'hardwired' pooling over an inappropriate large area in peripheral vision is a bottleneck for object identification, such feature integration incorporates statistical properties of the visual scene and enables fast texture segregation and segmentation.

In this thesis we are interested in the underlying 'hardwired' mechanisms of feature pooling in the early visual cortex. The hypothesis was that integration processes in response to differences of statistical properties in orientational texture-like stimuli induces the percept of an illusory contour. A method was developed that revealed the signature of integration processes at a fixed eccentricity using fMRI and a biologically plausible model for crowding.

In the following sections the signature obtained by the empirical (fMRI) and theoretical (virtual fMRI experiment using the crowding model) study is discussed. Furthermore, the comparison between both studies is done. Hence, the theoretical signature is fitted to the empirical data. Such comparison allowed the investigation of the underlying mechanisms that induced the illusory contour in

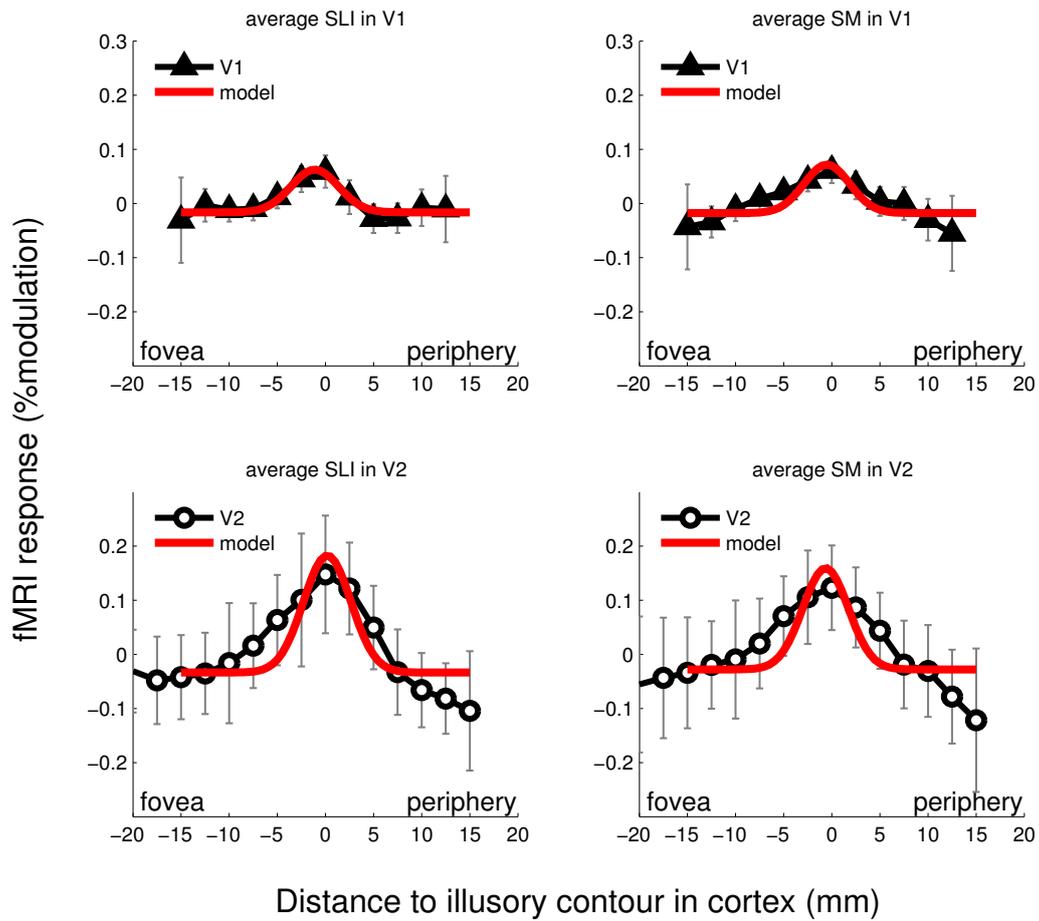
the empirical study. In other words, is the signature obtained with the fMRI study explained by 'hardwired' pooling of visual features as described by the crowding model?

## 6.1 Empirical Signature of Integration fields

The signature of processes attributed to integration fields was found using a sinusoid and block modulation of statistical differences between a disk and annulus. This revealed neural processes that correlated with the modulation of the statistical differences as function of cortical distance. To allow a robust comparison between the signatures obtained in the empirical and theoretical study the detrended projected amplitude signatures of the two modulation paradigms are averaged. This is possible because both the block and sinusoid modulation paradigm of the statistical differences between the disk and annulus correspond with the same BOLD modulation, i.e. the same fundamental frequency of 6 cycles per run is used. The black lines in figure 6.1 represent the averaged signatures for first-order (SLI) and second-order (SM) evoked illusory contours in V1 (diamond markers) and V2 (circle markers).

In the plotted signature of the detrended projected amplitude, a positive bell-shaped response is visible with the highest response centered at the disk and annulus transition, i.e. the location of the illusory contour in the visual cortex. Therefore, the conclusion that neural correlates of at least processes that are involved in the statistical analysis of the texture-like stimuli is prominent and is found as early as in V1. However, these signatures do not have to be attributed to processes in V1 and V2. There are many feedback connections from higher extrastriate cortical region that influences the responses in early visual cortical region (Dumoulin & Hess, 2006; Hupe et al., 1998). For instance, Dumoulin & Hess (2006) showed that V1 activity is inversely correlated with the structure in an image which is facilitated by processes in higher extrastriate regions.

One of the most profounding differences in the signatures is the higher amplitude of the bell-shaped curve in V2 compared with V1. It seems that V2 is more responsive to texture analysis induced by the texture-like stimuli. Kastner et al. (2000) corroborate with this finding. They showed in an fMRI study that V1 is less responsive, in contrast to higher visual cortical regions, to texture boundaries. Moreover, Mendola et al. (1999) showed in an fMRI study that responses to texture boundaries increased when moving up the hierarchy of visual processing. Therefore, supporting the notion that not a single area is involved in processing of texture boundaries. Thus, it would be interesting to investigate the signature of



**Figure 6.1:** Comparison Between fMRI and Theoretical study. The signature of the projected amplitude of the averaged condition SLI and SM are plotted on respectively the left and right side of the figure. The upper part of the figure is V1 and the lower part is V2. The black lines and markers are the signatures of the fMRI study whereas the red line is the fitted curve of the models sigma to the fMRI data. Hence, the amplitude and baseline was fitted according to a fitting procedure using the multidimensional unconstrained nonlinear minimization (Nelder-Mead) algorithm (Nelder & Mead, 1965).

extrastriate cortical region beyond V2 because feedback connections from higher cortical regions can influence early visual cortical regions and higher cortical regions seems to be more responsive to texture boundaries. Unfortunately, the scope and available time prohibited this research.

Another investigated condition is first- and second-order evoked illusory contours. Lateral inhibition (Blakemore et al., 1970), i.e. similar orientations in the vicinity of a cell's receptive field are inhibited, is a neural process that is ought to be important in pop-out and texture segregation (Kastner et al., 1997; Knierim & van Essen, 1992). Because lateral inhibition is only responsive to local changes we hypothesized that neural correlates of lateral inhibition shows only local responses to the condition of first-order evoked illusory contour. When comparing the empirical signature of the lateral inhibition (SLI) and mean (SM) condition in figure 6.1 a steeper bell-shaped curve can be read with good will in V1 and V2 (especially V1). However, more research is needed for stronger claims.

For strong claims, there are several important influences of this empirical research that should be taken into account. First, the fMRI research was performed on a small scale and it's main goal was to explore the possibilities of using texture-like stimuli where statistical differences induces the percept of an illusory contour. When data is presented individually at least 6 subjects need to be recorded to give prominent conclusions about the signature. Second, artifacts exist that influenced the results. For instance, uncontrolled eye movements during the experiment influences the correspondence between the retinotopic organization and the stimulus presentation. Furthermore, an intensity gradient artifact in the high detailed anatomical recording of the brain prohibits a clear gray and white matter boundary. This lead to wrong classification of voxels as gray matter. Lastly, the use of cortical distance as a measure of distance to the illusory contour is not flawless, i.e. cortical distance is not a one to one mapping of visual field distance.

## 6.2 Theoretical Signature of Integration Fields

The biologically model proposed by van den Berg et al. (2010) produces probability distributions of feature integration and hereby explains the crowding effect. Furthermore, the spatial extent of crowding in the visual field approached the size of integration fields. In this study we were interested in the segregation of a texture boundary using orientational texture-like stimuli and especially the extraction of an illusory contour. Therefore, a virtual fMRI experiment was performed. As such, the maximum amplitude of the net-responses of integration fields extracted with the crowding model were used as a measure for profoundness of angles con-

tained in the integration field. The virtual experiment resulted in a time series of profoundness of an integration field as function of cortical distance. Hence, a signature is found using the same experimental paradigm as the fMRI experiment.

In figure 5.3(b) the signature of the projected amplitude is plotted. Obviously, integration fields encapsulating only the disk or annulus are not responsive to the mean angle difference modulation between the disk and annulus. So, we only expect to see a amplitude modulation, corresponding with the stimulus modulation, in integration fields near the illusory contour. Furthermore, the strength of the modulation is stronger at the illusory contour because here the angle information is evenly divided into angle information originated from orientational patches of the disk and annulus. Thus, a positive bell-shaped curve is clearly visible in the signature.

Looking at the signatures obtained from the virtual experiment study there is a difference between the amplitude of the bell-shaped curve in the SLI and SM condition. This is understandable by the fact that in the SLI condition there is no noise. Hence, the net-responses of integration fields in the SLI condition are more profound and therefore has higher and steeper responses. Amplitude differences therefore increases in response to the same stimulus modulation. Note that the spatial extent of integration fields is not affected.

### 6.3 Comparison Empirical and Theoretical Signature

A comparison between the empirical and theoretical study provides an explanation of the found signature. To allow such comparison the signature obtained from the virtual experiment of the crowding model was fitted to the empirical study. This is done by fitting the theoretical signature to a Gaussian distribution using the multidimensional unconstrained nonlinear minimization (Nelder-Mead) algorithm (Nelder & Mead, 1965). The sigma of this fit is used to fit a Gaussian distribution to the fMRI signature according to the following Gaussian function

$$f(x; \mu, \sigma) = base + A \left( \frac{\sqrt{2\pi}}{\sigma_{model}} \cdot \exp \left( -\frac{(x - x_0)^2}{2\sigma_{model}^2} \right) \right)$$

where *base* is a baseline of activation (base of signature is not at amplitude 0 in theoretical study), *A* the amplitude of the distribution,  $x_0$  the center of the distribu-

tion and  $\sigma_{model}$  the variance of the distribution that is determined by the theoretical fit to a Gaussian distribution. So, only the amplitude and baseline of activation is fitted to the fMRI data using the Nelder-Mead algorithm and  $\sigma_{model}$  is determined by the fit to the theoretical signature. The center of the fit, i.e.  $x_0$ , is fixed at the illusory contour, that is 0 millimeter distance. In figure 6.1 the red lines are the fitted curves to the fMRI signatures for every condition and visual cortical regions V1 and V2. Note that the cortical mapping in the theoretical study is only modeled for V1. Hence, the same cortical mapping is used for the theoretical signature obtained for V2.

There seems to be a reasonable fit of the theoretical and empirical signatures. This is an indication that similar processes are found. However, the empirical signature shows a much wider bell-shaped curve. Several theories could explain this discrepancy. First, it could be plausible that cells in the visual cortex exist that process texture information on a higher level than compulsory averaging of textures in integration fields. Hence, cells that combine net-responses of the integration fields or cells that encode combinations of orientation. The latter is a mechanism already found in V2 (Anzai et al., 2007). Second, individual differences of the critical region alters the width of the bell-shaped curve (approximated by  $\sigma_{model}$ ). Third, as already mentioned, higher cortical regions influence cortical responses in the early visual cortex. Thus, the shape of the signature could be transformed due to higher cortical area influences.

Last remark about the comparison between the fitted theoretical and empirical signature, is that when looking at the signatures obtained in the theoretical study there are no differences between the width of the bell-shaped curve in the SLI and SM condition. The reason for this is that lateral inhibition is not modeled in the crowding model and therefore such comparison, between the empirical and theoretical experiment, is redundant. Furthermore, there is no amplitude difference in the SLI and SM condition in the empirical study in contrast to the theoretical study. This is obvious considering that fMRI is an indirect measure of neural activity (ratio between oxygenated and deoxygenated hemoglobin) and the theoretical study models neural spike rates. Hence, there is no fixed mapping between the amplitude of both studies.

## 6.4 General discussion

Human vision is very complicated and sophisticated. This thesis focused on only a small part of the underlying mechanisms of human vision and especially the crowding effect. Crowding is getting lots of attention the last few decades and

seems to be very important in texture analysis and the limitations of object recognition (Levi, 2008). Investigating the underlying mechanisms of crowding increases our understanding of human vision and can be applied to different fields of research. Especially, the field of artificial intelligence is interested in 'smart' solutions that are closely related or even mimic processes of human vision. For instance, a better understanding of texture analysis in human brains can attribute to robust and reliable texture analysis in computer vision. Furthermore, the field of neuro-ergonomics lack research that explains underlying mechanisms in the brain that could contribute to better ergonomic designs. For instance, the spatial extent wherein crowding takes place could be used as a layout constraint in high demanding designs where lots of information need to be presented simultaneously.

In further research, it is very interesting to investigate the signature of illusory contours induced at other eccentricities. This allows the comparison of the spatial extent of integration processes. Hence, perhaps the spatial extent of the critical region can be linked to the size of the bell-shaped signatures. If these integration processes can be linked, we hypothesize that it is possible to fit cortical responses of a moving illusory border to a harmonic function. Thus, a new method of standard visual field mapping, using the population receptive field (pRF) model analysis, can be performed; estimating the size of integration fields. Here, stimuli can consist of a textured background and textured rotating wedges or expanding rings. The edge of the wedge or ring is illusory. That is, only difference in mean orientation separates 'figure' from 'ground'.

The problem with integration field size mapping is that the illusory contour near the edge of the stimulus presentation will not yield satisfactory results because outside the annulus no orientational stimuli can be presented. In other words, when the spatial extent of an integration field encapsulates the stimulus presentation outside the annulus will impose a problem. However, newly developed techniques use MR compatible video goggles and enable stimulus presentation in the full visual field of a subject (Engström et al., 2005).

To summarize, the exploratory research outlined in this thesis, in search of texture integration in the early visual cortex, tried to investigate feature integration processes that are important in texture analysis and developed new methods to do so. Further research is necessary to reveal the true nature of the fMRI signal in response to the developed texture-like stimuli and could contribute to a better understanding of human vision and especially crowding. Therefore, exploring the capabilities and constraints of texture analysis of visual scenes in artificial and human vision.

## References

- Anzai, A., Peng, X., & Van Essen, D. C. (2007). Neurons in monkey visual area v2 encode combinations of orientations. *Nature Neuroscience*, 10, 1313 – 1321.
- Bhatt, R., Carpenter, G. A., & Grossberg, S. (2007). Texture segregation by visual cortex: Perceptual grouping, attention, and learning. *Vision Research*, 47(25), 3173–3211.
- Blakemore, C., Carpenter, R. H. S., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature*, 228, 237–260.
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226, 177–178.
- Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, 13, 767–782.
- Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human v1. *Journal of Neuroscience*, 16, 4207– 4221.
- Boynton, G. M. & Finney, E. M. (2003). Orientation-specific adaptation in human visual cortex. *Journal of Neuroscience*, 23, 8781–8787.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.*, 10, 433–436.
- Chung, S. T. L., Li, R. W., & Levi, D. M. (2007). Crowding between first- and second-order letter stimuli in normal foveal and peripheral vision. *Journal of Vision*, 7(2):10, 1–13.

- Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., & Marchal, G. (1995). Automated multi-modality image registration based on information theory. *Information Processing in Medical Imaging*, (pp. 263–274).
- Cornelissen, F. W., Wade, A. R., Vladusich, T., Dougherty, R. F., & Wandell, B. A. (2006). No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex. *Journal of Neuroscience*, 26, 3634–3641.
- Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human photoreceptor topography. *Journal of Comparative Neurology*, 292, 497–523.
- Dakin, S. C., Bex, P. J., Cass, J. R., & Watt, R. J. (2009). Dissociable effects of attention and crowding on orientation averaging. *Journal of Vision*, 9(11):28, 1–16.
- DeYoe, E. A., Carman, G., Bandettini, P., Glickman, S., Cox, R., Miller, D., & Neitz, J. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc National Academy of Sciences*, 93, 2382–2386.
- Dumoulin, S. O. & Hess, R. F. (2006). Modulation of v1 activity by shape: Image-statistics or shape-based perception? *Journal of Neurophysiology*, 95(6), 3654 – 3664.
- Dumoulin, S. O. & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2), 647–660.
- Engel, S. A., G. H., G., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional mri. *Cereb Cortex*, 7, 181–192.
- Engel, S. A., Rumelhart, D. E., Wandell, B. A., Lee, A. T., Glover, G. H., Chichilnisky, E. J., & Shadlen, M. N. (1994). fmri of human visual cortex. *Nature*, 369, 525.
- Engström, M., Ragnehed, M., & Lundberg, P. (2005). Projection screen or video goggles as stimulus modality in functional magnetic resonance imaging. *Magnetic Resonance Imaging*, 23(5), 695–699.
- Fang, F. & He, S. (2008). Crowding alters the spatial distribution of attention modulation in human primary visual cortex. *Journal of Vision*, 8(9):6, 1–9.

- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vision Research*, 33, 173–193.
- Furmanski, C. & Engel, A. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience*, 3, 535–536.
- Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2002). *Cognitive Neuroscience, Second Edition*, chapter 5, (pp. 148–192). W.W. Norton & Company, Inc.
- Grigorescu, C., Petkov, N., & Westenberg, M. A. (2003). Contour detection based on non-classical receptive field inhibition. *IEEE Transactions on Image Processing*, 12(7), 729–739.
- Grill-Spector, K. & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, 7, 649–677.
- Harel, N., Lee, S. P., Nagaoka, T., Kim, D. S., & Kim, S. G. (2002). Origin of negative blood oxygenation level-dependent fmri signals. *Journal of Cerebral Blood Flow Metabolism*, 22, 908–917.
- Hubel, D. H. & Wiesel, T. N. (1977). Ferrier lecture: Functional architecture of macaque monkey visual cortex. In *Royal Society of London. Series B*, volume 198 (pp. 1–59).
- Hupe, J. M., James, A. C., Payne, B. R., Lober, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by v1, v2 and v3 neurons. *Nature*, 394, 784–787.
- Intriligator, J. & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, 43, 171–216.
- Jezzard, P., Matthews, P. M., & Smith, S. M. (2001). *Functional MRI: An Introduction to Methods*. Oxford University Press, USA.
- Jones, J. P., Stepnoski, A., & Palmer, L. A. (1987). The two-dimensional spectral structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58, 1212–1232.
- Julesz, B. (1975). Experiments in the visual perception of textures. *Scientific American*, 232, 34–43.

- Kastner, S., de Weerd, P., & Ungerleider, L. G. (2000). Texture segregation in the human visual cortex: a function mri study. *Journal of Neurophysiology*, 83, 2453–2457.
- Kastner, S., Nothdurft, H. C., & Pigarev, I. N. (1997). Neuronal correlates of pop-out in cat striate cortex. *Vision Research*, 37, 371–376.
- Knierim, J. J. & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area v1 of the alert macaque monkey. *Journal of Neurophysiology*, 67(4), 961–980.
- Korte, W. (1923). Uber die gestaltauffassung im indirekten sehen. *Zeitschrift fur Psychologie*, 93, 17–82.
- Kouhsari, L. M. & Rajimehr, R. (2005). Subliminal attentional modulation in crowding condition. *Vision Research*, 45(7), 839–844.
- Lee, T. S. (1996). Image representation using 2d gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine intelligence*, 18(10), 959–971.
- Levi, D. M. (2008). Crowding - an essential bottleneck for object recognition: a mini-review. *Vision Research*, 48, 635–654.
- Levi, D. M., Song, S., & Pelli, D. G. (2007). Amblyopic reading is crowded. *Journal of Vision*, 7(2):21, 1–17.
- Liu, T., Jiang, Y., Sun, X., & He, S. (2009). Reduction of the crowding effect in spatially adjacent but cortically remote visual stimuli. *Current biology*, 19(2), 127–32.
- Logothetis, N. & Wandell, B. (2004). Interpreting the bold signal. *Annual Review of Physiology*, 66, 735–769.
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fmri signal. *Nature*, 412, 150–157.
- Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *Journal of Vision*, 9(4):14, 1–18.
- Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. H. (1999). The representation of illusory and real contours in human cortical visual areas revealed by fmri. *Journal of Neuroscience*, 19, 8560–8572.

- Motter, B. C. & Simoni, D. A. (2007). The roles of cortical image separation and size in active visual search performance. *Journal of Vision*, 7(2):6, 1–15.
- Nelder, J. A. & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4), 308–313.
- Nestares, O. & Heeger, D. J. (2000). Robust multiresolution alignment of mri brain volumes. *Magn Res Med*, 43, 705–715.
- Oliva, A., Mack, M. L., Shrestha, M., & Peeper, A. (2004). Identifying the perceptual dimensions of visual complexity of scenes. In *The 26th Annual Meeting of the Cognitive Science Society* Chicago.
- Parkes, L., Lund, J., Angelucci, A., & Solomon, J. A. and Morgan, M. J. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4, 739–744.
- Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.*, 10, 437–442.
- Pelli, D. G. (2008). Crowding: a cortical constraint on object recognition. *Current Opinion Neurobiology*, 18, 445–451.
- Pelli, D. G., Majaj, N. J., Raizman, N., Christian, C. J., Kim, E., & Palomares, M. C. (2009). Grouping in object recognition: The role of a gestalt law in letter identification. *Cognitive Neuropsychology*, 26(1), 36–49.
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, 4(12):12, 1136–1169.
- Pelli, D. G. & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 11(10), 1129–1135.
- Petrov, Y. & Popple, A. V. (2007). Crowding is directed to the fovea and preserves only feature contrast. *Journal of Vision*, 7(2):8, 1–9.
- Pouget, A., Zemel, R. S., & Dayan, P. (2000). Information processing with population codes. *Nature*, 1(2), 125–132.
- Pöder, E. (2006). Crowding, feature integration, and two kinds of "attention". *Journal of Vision*, 6(2):7, 163–169.

- Renninger, L. W. & Malik, J. (2004). When is scene identification just texture recognition? *Vision Research*, 44, 2301–2311.
- Roy, C. S. & Sherrington, C. S. (1890). On the regulation of the blood-supply of the brain. *The journal of physiology*, 11, 85–108.
- Schira, M., Wade, R., & Tyler, C. (2007). The two-dimensional mapping of the central and parafoveal visual field to human visual cortex. *Journal of Neurophysiology*, 97, 4284–95.
- Schwartz, E. L. (1980). Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. *Vision Research*, 20(8), 645–669.
- Schyns, P. G. & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5, 195–200.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., & Tootell, R. B. H. (1995). Borders of multiple visual areas in human revealed by functional magnetic resonance imaging. *Science*, 268, 889–893.
- Shmuel, A., Augath, M., Oeltermann, A., & Logothetis, N. K. (2006). Negative functional mri response correlates with decreases in neuronal activity in monkey visual area v1. *Nature Neuroscience*, 9, 569–577.
- Smith, M. A. (2006). Surround suppression in the early visual system. *The Journal of Neuroscience*, 26(14), 3624–3625.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23(S1), 208–219.
- Talairach, J. & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Thieme Medical Publishers. New York.
- Teo, P. C., G., S., & Wandell, B. A. (1997). Creating connected representations of cortical gray matter for functional mri visualization. *IEEE Med. Transactions*, 16 (6), 852–863.

- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.
- Toet, A. & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32, 1349–1357.
- Tootell, R. B. H., Hadjikhani, N. K., Mendola, J. D., Marret, S., & Dale, A. M. (1998). From retinotopy to recognition: fmri in human visual cortex. *Trends in Neuroscience*, 2(5), 174–183.
- van den Berg, R., Cornelissen, F. W., & Roerdink, J. B. T. M. (2009). A crowding model of visual clutter. *Journal of Vision*, 9(4):24, 1–11.
- van den Berg, R., Roerdink, J. B. T. M., & Cornelissen, F. W. (2007). On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. *Journal of Vision*, 7(2), 1–11.
- van den Berg, R., Roerdink, J. B. T. M., & Cornelissen, F. W. (2010). A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Computational Biology*, 6(1), e1000646. Submitted for publication.
- Wandell, B. A. (1999). Computational neuroimaging of human visual cortex. *Annual Review of Neuroscience*, 10(22), 145–173.
- Wandell, B. A., Chial, S., & Backus, B. (2000). Visualization and measurement of the cortical surface. *Journal of Cognitive Neuroscience*, 12(5), 739–52.
- Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, 56(2), 366–383.
- Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., & Smith, S. M. (2009). Bayesian analysis of neuroimaging data in fsl. *NeuroImage*, 45, 173–186.
- Woolsey, T. A., Rovainen, C. M., Cox, S. B., Henegar, M. H., Liang, G. E., Liu, D., Moskalenko, Y. E., J., S., & Wei, L. (1996). Neuronal units linked to microvascular modules in cerebral cortex: response elements for imaging the brain. *Cereb Cortex*, 6, 647–660.
- Xing, J. & Heeger, D. J. (2000). Center-surround interactions in foveal and peripheral vision. *Vision Research*, 40, 3065–3072.

- Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006). User-guided 3d active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3), 1116–1128.
- Zemel, R. S., Dayan, P., & Pouget, A. (1998). Probabilistic interpretation of population codes. *Neural Computation*, 10(2), 403–430.

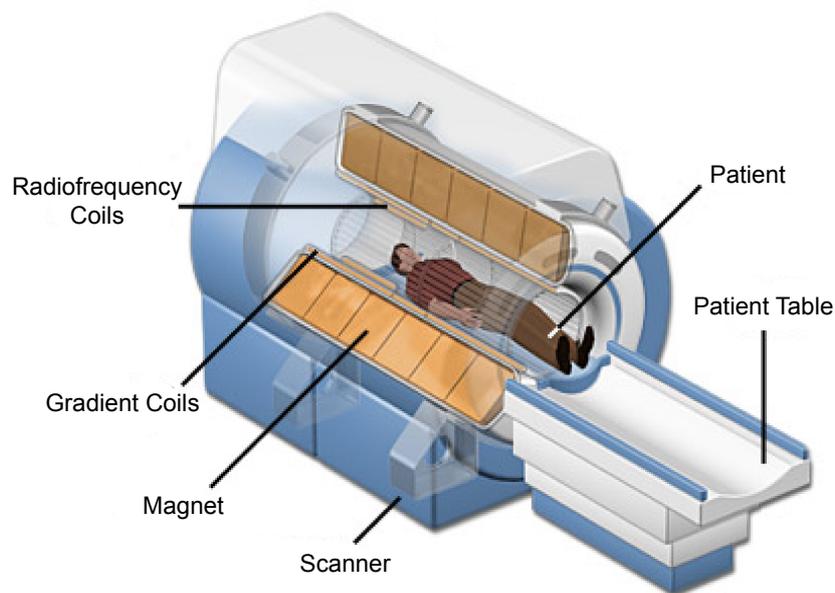
## Brain Imaging

The imaging technique, to measure brain activity, used in this thesis is a specialized technique of magnetic resonance imaging (MRI), i.e. functional MRI (fMRI). The (f)MRI technique is used extensively because it's wide availability, relative low-costs and non-invasiveness. MRI is used for anatomical imaging purposes that enables the diagnosis of tumors and other abnormalities in the human body. Nowadays, neuroscientists apply the fMRI technique to investigate the involvement of brain structures in cognitive functional processes and their underlying mechanisms. The principles behind the (f)MRI technique are outlined in this chapter. See Jezzard et al. (2001) for a more thoroughly explanation of the technique and methods in (f)MRI.

### **A.1 The (f)MRI technique**

An MRI scanner is constructed of a very powerful magnet. Patients, lying on a movable table bed, are shoved inside a tube that is surrounded by the magnet, see figure A.1. MRI scanners reach magnetic field strengths of 7 Tesla. In contrast, earth's magnetic field is more than 20.000 times weaker than the magnetic field produced by MRI scanners.

The MRI technique is based on nuclear magnetic resonance caused by isotopes with an uneven number of core particles. These isotopes have a charged core and therefore have a quantum mechanic property of spin that causes a small magnetic field. The strong magnetic field induced by the MRI magnet results in the alignment of the magnetic spin to the direction of the magnetic field. The magnetic spin aligns in two possible energy states, the so called 'eigenstates'. If an electromagnetic radiofrequency pulse with the right amount of energy is send to



**Figure A.1: MRI Scanner Setup.** This figure shows an MRI scanner setup. Patient is moved inside the MRI tube. A strong magnet causes isotopes to orient to the magnetic field. The alignment of the isotopes is disturbed by sending radio frequency pulses. When the radiofrequency pulse is stopped, the alignment of the isotopes return to the orientation of the magnetic field. This is accompanied by the emission of energy (photons). The scanner records the emitted photons which is a measure of the isotope density. Adding a gradient of the magnetic field allows to measure densities at specific locations. (source: <http://www.ehow.com>)

these atoms, the absorption of the radiofrequencies energy will flip the alignment. The reason for this flip is because radiofrequencies behave as an external oscillating magnetic field that causes resonant absorption of the energy in protons that are sensitive to that specific frequency. When the radiofrequency pulse is stopped, the magnetic field of the isotopes will return to its original state, the so called relaxation period. The relaxation period is accompanied with a release of energy, i.e. the emission of photons. Such emission of photons is recorded in the MRI scanner. When the longitudinal magnetization recovery of the spin is measured, we speak of a T1-weighted scan.

By adding a gradient in the magnetic field, produced by gradient coils inside the tube of the MRI scanner, different wavelengths of the restoring spin are measured because a stronger magnetic field causes faster relaxation periods. Consequently, the different wavelengths emitted in the relaxation period reveal the origin of the emitted photons in space. The smallest unit of a location measurement is called a voxel. The wavelength of the emitted photons of the investigated isotope is recorded for every voxel with a specified location and dimension. A contrast map can be constructed from the measured photon emission in the voxel grid, where the amount of emitted energy is an indirect measure of density differences of a particular isotope in space.

### *A.1.1 Anatomical Imaging*

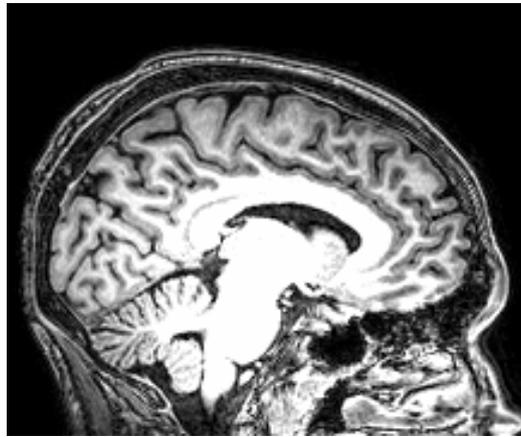
A widely investigated isotope is hydrogen. Structures in the human body are composed of tissues with different densities of hydrogen. Imaging these densities will enable researchers to distinguish different structures in the human body. Moreover, lesions, tumors and strokes can be diagnosed. The density of hydrogen is measured when a radiofrequency, with the right amount of energy<sup>1</sup>, flips the alignment of the magnetic spin of hydrogen isotopes. As a result, a contrast map can be constructed to provide a visualization of relative hydrogen densities. See figure A.2 for a sagittal slice of a T1-weighted anatomical scan of the brain. Densities of hydrogen molecules are colored on a scale from black to white scaling respectively low and high hydrogen densities.

### *A.1.2 Functional Imaging*

Changes in blood flow in the human brain are known to be closely linked to neural activity (Roy & Sherrington, 1890; Logothetis et al., 2001). Especially, in red

---

<sup>1</sup>The so called Larmor frequency is the effective radiofrequency for flipping the alignment of the magnetic spin of hydrogen.



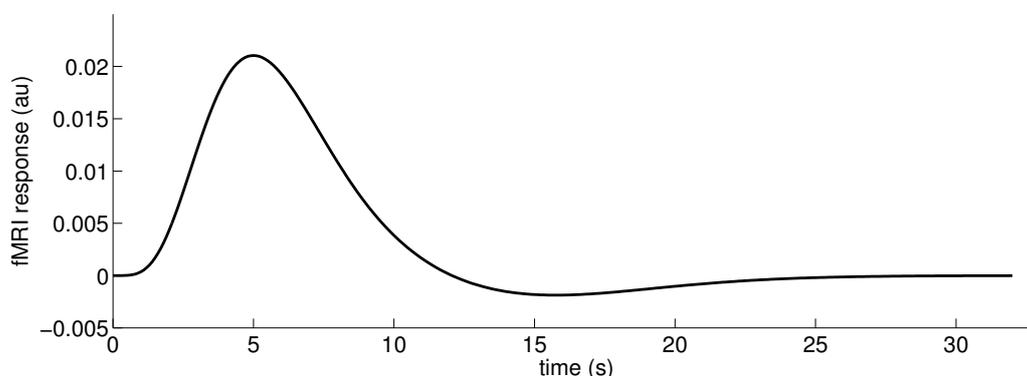
**Figure A.2:** *Hydrogen Contrast of Brain. Sagittal slice of a T1-weighted MRI scan of the brain.*

blood cells a change of the protein hemoglobin occurs. Hemoglobin is a protein that is involved in the transportation of oxygen in the human body. Hemoglobin is oxygenated in the lungs and carries the oxygen to cells in the human body through the cardiovascular system. When cells require oxygen hemoglobin is deoxygenated. These two different ‘states’ of hemoglobin can be independently imaged because oxygenated and deoxygenated hemoglobin have different magnetic resonance properties. Therefore, oxygenated and deoxygenated hemoglobin are sensitive to a dissimilar radiofrequency. The proportion of oxygenated and deoxygenated hemoglobin is known as the Blood Oxygenation Level Dependent (BOLD) signal. Higher BOLD-signals resemble a high concentration of oxygenated, as opposed to deoxygenated, hemoglobin.

Neuronal activity is accompanied by extracting oxygen from the red blood cells. The drop of oxygenated hemoglobin is compensated by an increased cerebral blood flow (CBF). The increased CBF and oxygen intake of active structure changes the BOLD-signal. The change of the BOLD-signal, after increased oxygen demands in human brain structures, follows approximately the hemodynamic response function (HRF), see figure A.3. The HRF rises to a peak after 6-10 seconds<sup>2</sup> and will then drop below baseline. Hereafter, the HRF rises gradually to baseline and has died out after approximately 30 seconds. Voxels that shows a similar time course as the HRF indicates neural activity within those voxel.

---

<sup>2</sup>One would expect an initial drop in the BOLD signal because cells extract oxygen from red blood cells. But, the CBF overcompensates for the drop in oxygenated hemoglobin. Hence, the HRF initially shows a rise in BOLD response. Though, a very small drop in the HRF can be measured with very sensitive imaging protocols.

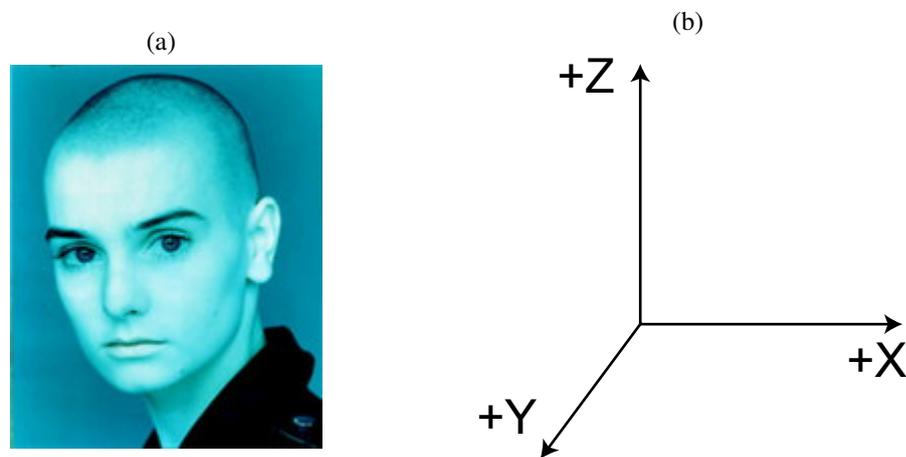


**Figure A.3:** *Hemodynamic Response Function. A two-gamma modeled HRF. The HRF is the time-course of the BOLD-signal in response to neural activity. The signal modulation in arbitrary units is plotted as function of time.*

## A.2 (f)MRI Procedure

The scanning procedure in MRI research depend on adjustable parameters and standards. The location and dimension of the recorded voxels are determined by the experimenter and scanner operator. The default format of the recorded voxels is the radiological convention. The radiological convention format is;  $+X$  is Left,  $+Y$  is Anterior and  $+Z$  is Superior (LAS). Hence, this format intuitively uses the point of view from the subject. However, when analysing (f)MRI data the neurological convention is more intuitively:  $+X$  is Right,  $+Y$  is Anterior and  $+Z$  is Superior (RAS), where only the left and right orientation is swapped. See figure A.4 for the graphical illustration of the neurological convention.

Another important parameter is repetition time (TR). Measurements of relaxation periods for every voxel is not instant and therefore the TR determines the duration of a single scan. A larger TR allows higher resolutions but prohibit ‘fast’ alternating stimuli. It is common that when measuring the BOLD-response the TR is  $\geq 1500$  milliseconds and depends on the trade-off for temporal and spatial resolution. Hence, high temporal resolution and low spatial resolution results in a low TR and vice versa.



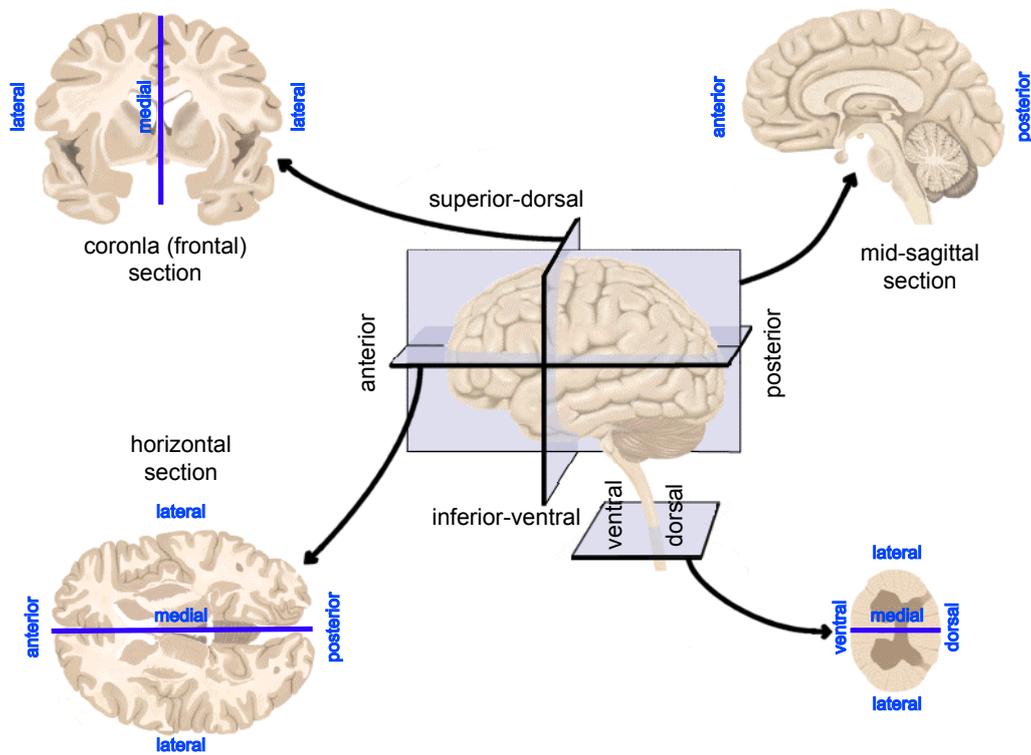
**Figure A.4:** *Neurological Convention.* (a) The anatomy of the human brain is more intuitively grasped with (b) the neurological convention when analysing (f)MRI data. Here, the direction of the voxels with regard to the subject are determined as follows; +X is Right +Y is Anterior +Z is Superior (RAS).

# Appendix **B**

## Anatomical Terminology

A graphical representation of the anatomical directions is shown in figure B.1. For easy reference the brain is viewed in three two-dimensional slices for each direction, i.e. coronal, sagittal and horizontal sections. The front and the back of the brain is respectively referred to anterior and posterior. The line that splits the right and left hemisphere is the medial line. Therefore, moving towards this line is referred to medial cortical regions. The outer surface of the brain is referred to lateral. Furthermore, contralateral and ipsilateral is the terminology to refer to respectively the other side or same side. Hence, the left visual field is processed in the contralateral (right) hemisphere.

The dorsal and ventral direction is respectively references to the top and bottom of the brain. Note that in figure B.1 ventral and dorsal are also drawn for the brainstem. However, these directions do not resemble a top and bottom discrepancy because the anatomical directions are first introduced for animals. Hence, the spinal column of animals, for instance a dog, is stretched in horizontal position.



**Figure B.1: Anatomical Directions.** The important anatomical directions of the human brain. A coronal, sagittal and horizontal section is shown with corresponding locations and directions. (source: <http://jdjacobson.net>)

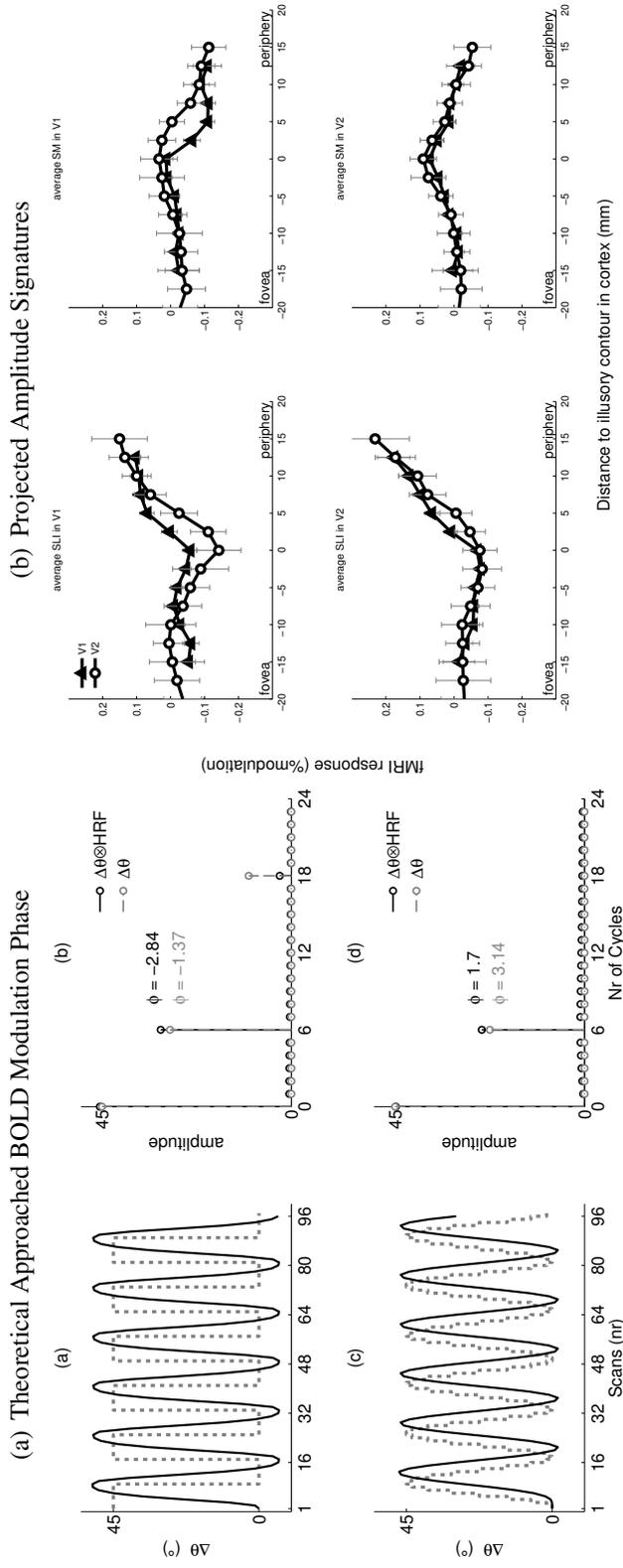
## Supplementary

### C.1 Projected Amplitude

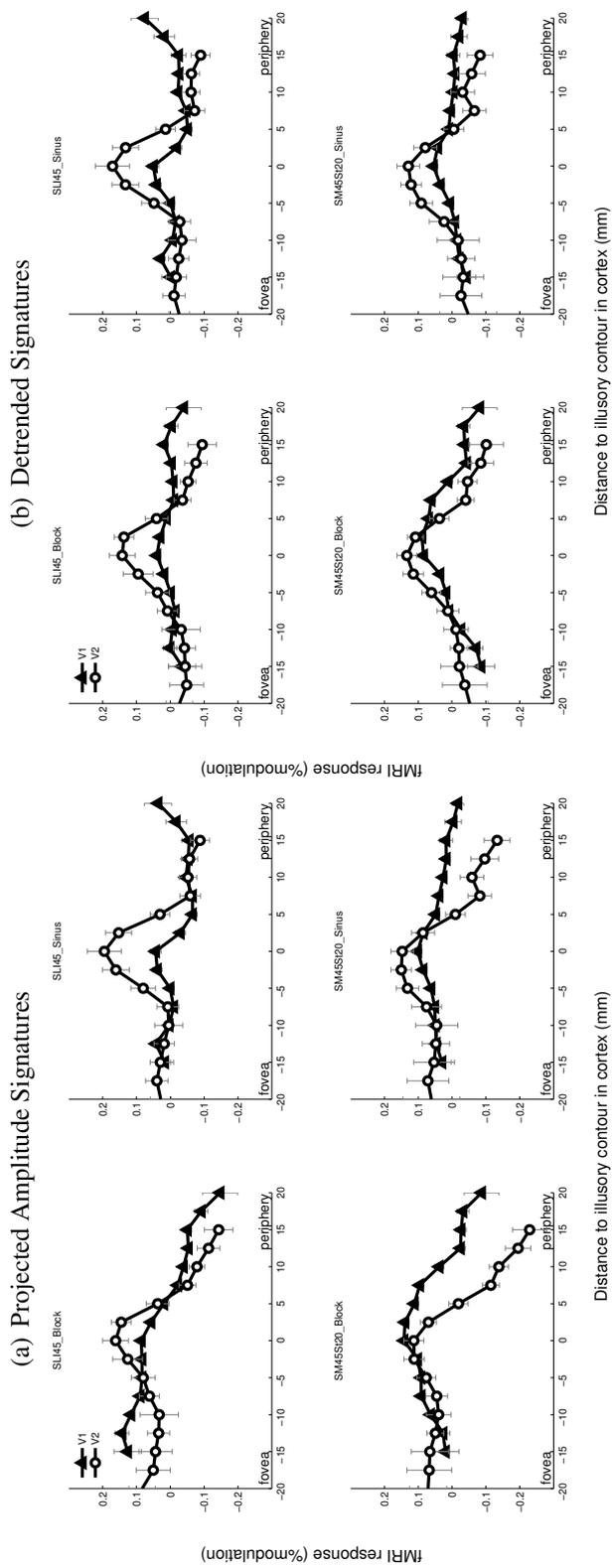
The BOLD modulation of the time series is out of phase with the stimulus modulation. This is obvious considering the delay of the peak response of the HRF. See figure C.1 for an analysis of the expected time series and amplitude spectrum of the stimulus and BOLD modulation. Here, the expected phase delay of the measured BOLD modulation in the block and sinusoid stimulus condition are respectively  $-2.84$  rad and  $1.7$  rad. So, these values are used as the synchronization phase to assure temporal correspondence between the stimulus and BOLD modulation. See figure C.1 for the projected amplitude. Although the signature shows a stronger peak at the illusory contour representation, the peak in the block and sinusoid modulation are not consistent. We would expect no difference in the sign of the peak of the modulation paradigms. Thus, the theoretical approached phase of the stimulus modulation and HRF used in the projected amplitude is not satisfactory.

### C.2 Individual Data

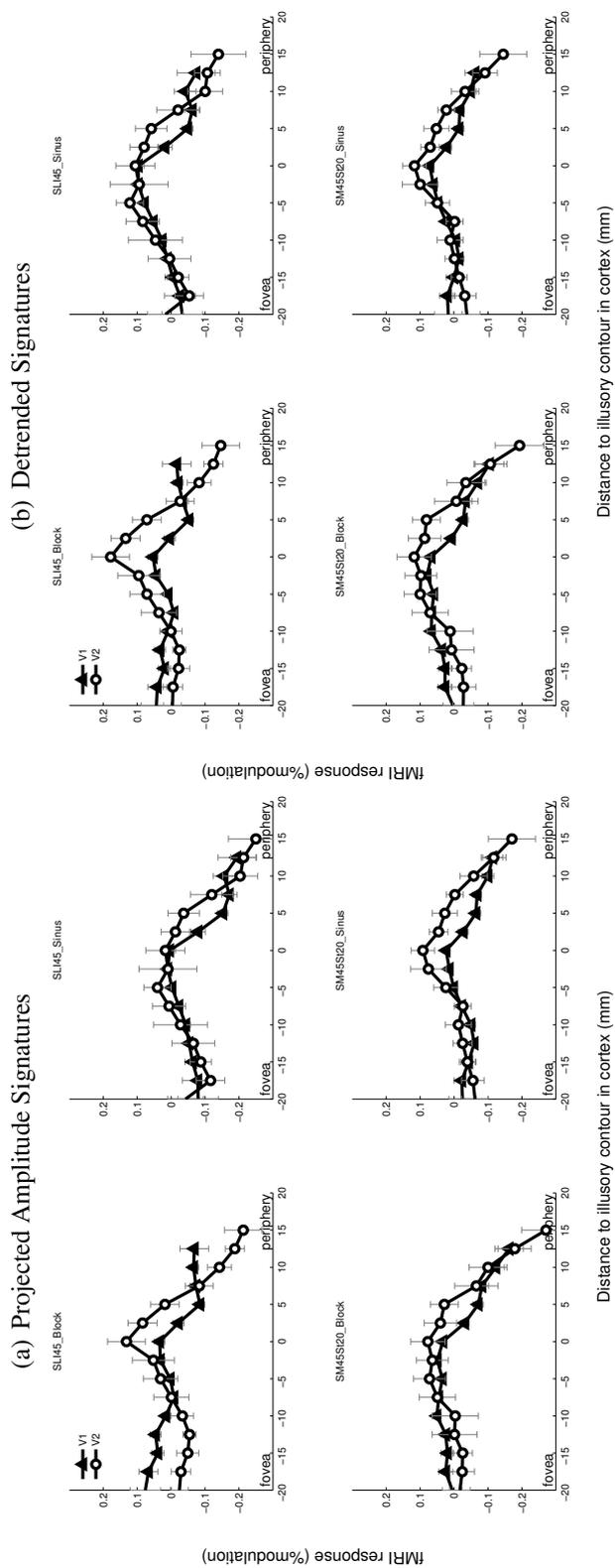
Signatures of individual data is shown in figure C.2 and C.3 for respectively subject KH and SF. The projected amplitude is the amplitude synchronized with the modulation phase at the illusory contour. The detrended projected amplitude are plotted as well to increase visibility and allow for a easier comparison.



**Figure C.1:** Expected Phase of BOLD Modulation and Resulting Projected Amplitude Signature. The left figure (a) wherein (a) the dotted gray line is the mean orientation modulation,  $\Delta\theta$ , according to the block paradigm and the solid black line is the stimulus modulation convolved with the HRF, i.e. the expected BOLD modulation. (b) The single-sided amplitude spectrum of the stimulus and BOLD modulation, where the phase of the fundamental frequency (6 cycles per run) is shown in radians. (c) The sinusoid experimental paradigm is the dotted gray line. Convolution of the sinusoid and the HRF is the black solid line. (d) In the Fourier spectrum, there is a peak at the fundamental frequency and the phase to the stimulus and BOLD modulation is shown in radians. Note that the block modulation has a phase delay as opposed to the sinusoid modulation. Therefore, the phase in the spectra differ. (b) Here, the signatures of the projected amplitude in V1 and V2 are plotted for every condition and stimulus modulation paradigm. The synchronization phase is the theoretical approached phase delay. The projected amplitude corrects the amplitude and assures a fixed temporal correspondence between the expected BOLD modulation and the measured BOLD modulation. Note that the signatures are not detrended to preserve the most raw data.



**Figure C.2:** Signatures of Projected Amplitude and Detrended Projected Amplitude of Subject KH. In all figures the signatures were analyzed in V1 and V2 voxels and are plotted for every condition. In (a) the projected amplitude and in (b) the detrended projected amplitude is plotted as function of cortical distance to the illusory contour.



**Figure C.3:** Signatures of Projected Amplitude and Detrended Projected Amplitude of Subject SF. In all figures the signatures were analyzed in V1 and V2 voxels and are plotted for every condition. In (a) the projected amplitude and in (b) the detrended projected amplitude is plotted as function of cortical distance to the illusory contour.