

# Mechanisms of surface perception uncovered by visual illusions

Citation for published version (APA):

Jans, B. (2011). *Mechanisms of surface perception uncovered by visual illusions*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. <https://doi.org/10.26481/dis.20111104bj>

**Document status and date:**

Published: 01/01/2011

**DOI:**

[10.26481/dis.20111104bj](https://doi.org/10.26481/dis.20111104bj)

**Document Version:**

Publisher's PDF, also known as Version of record

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# **Mechanisms of Surface Perception Uncovered by Visual Illusions**

Bert Jans

August 29, 2011



# Mechanisms of Surface Perception Uncovered by Visual Illusions

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,  
op gezag van de Rector Magnificus, Prof. mr. G. P. M. F. Mols  
volgens het besluit van het College van Decanen,  
in het openbaar te verdedigen op vrijdag 4 november 2011 om 12u.

door

Bert Jans

**Promotor: Prof. Dr. Peter De Weerd**

**Copromotor: Prof. Dr. Rainer Goebel**

**Beoordelingscommissie:**

**Prof. Dr. Bernadette Jansma (voorzitter)**

**Prof. Dr. Leo Blomert**

**Prof. Dr. Wim Vanduffel (Katholieke Universiteit, Leuven, België)**

**Dr. Elia Formisano**

**Dr. Alex Sack**

**The work in this thesis was supported by grants from the Netherlands Organization of Scientific Research (NWO) open competition (400-04-036), VENI (451-07-014) and VICI (453-04-002).**

# Contents

<b>1 Introduction</b>	<b>9</b>
Visual illusions as a tool to study mechanisms of surface perception . . . . .	10
Boundary and surface representations: Neurophysiology . . . . .	13
Computational Models . . . . .	14
Research question and paradigms . . . . .	15
Surface brightness perception studied via an induction paradigm	16
Texture surface perception studied via a Troxler paradigm . .	18
Methodological considerations . . . . .	20
Thesis overview . . . . .	21
References . . . . .	22
<b>2 Calculating group level activity landscapes of spatial fMRI data</b>	<b>29</b>
Abstract . . . . .	30
Introduction . . . . .	30
Methods . . . . .	32
fMRI experiments . . . . .	32
Cortex based alignment (CBA) . . . . .	33
Functional localiser analysis and experimental data . . . . .	34
Simulations . . . . .	34
SNR calculation . . . . .	34
Results . . . . .	34
CBA versus PBA . . . . .	35
PBA algorithm and illustration . . . . .	35
Computing a mesh to mesh mapping . . . . .	38
Functional alignment of a localizer experiment: comparison between PBA and CBA . . . . .	40

## *Contents*

Functional alignment of simulated data: comparison between PBA and CBA . . . . .	43
Discussion . . . . .	46
Functional alignment in topographic maps using PBA . . . . .	46
Applications of PBA in high field imaging . . . . .	50
Conclusion . . . . .	50
References . . . . .	51
<b>3 A correlate of brightness induction in human V2</b>	<b>57</b>
Abstract . . . . .	58
Introduction . . . . .	58
Methods . . . . .	59
Stimulus Design . . . . .	60
Psychophysical Ratings of Induction . . . . .	63
Functional Imaging Experimental Design . . . . .	64
Scanning Parameters . . . . .	65
Preprocessing and Analysis of MR Images . . . . .	65
Retinotopic Mapping and Functional Alignment . . . . .	66
Results . . . . .	67
Psychophysics . . . . .	67
Neural Correlates of Brightness Induction on the Cortical Surface . . . . .	67
Control Experiments . . . . .	71
Association Between Ratings and Neural Correlate of Brightness Illusion . . . . .	73
Discussion . . . . .	73
References . . . . .	77
<b>4 Induced brightness filling-in in early visual areas: neural modeling of fMRI data in a common brain space</b>	<b>81</b>
Abstract . . . . .	82
Introduction . . . . .	82
Methods . . . . .	85
Brightness illusion and brightness control design . . . . .	85
Large-scale neural network model . . . . .	85

fMRI data . . . . .	89
fMRI and neural net integration in Common Brain Space framework . . . . .	89
Results . . . . .	90
Activity in M-<retina> and M-< $V1_{blob}$ > reflect physical and perceived brightness respectively . . . . .	94
Activity in M-< $V1_{blob}$ > corresponds to hemodynamic correlates of perceived brightness in V1 . . . . .	96
Discussion . . . . .	101
Neural mechanisms of brightness perception . . . . .	101
Our model and other computational models of brightness perception . . . . .	104
Large-scale neural models and neuroimaging data in a common brain space: Current possibilities and future perspectives . . . . .	109
References . . . . .	111
<b>5 A spatio-temporal correlate of texture filling-in in human V2</b>	<b>119</b>
Abstract . . . . .	120
Introduction . . . . .	120
Methods . . . . .	122
Retinotopic mapping and localizer stimuli . . . . .	122
Experimental design and participant instructions . . . . .	123
Apparatus . . . . .	124
Functional data preprocessing . . . . .	124
Functional and psychophysical data analysis . . . . .	125
Eye movement data . . . . .	126
Results and discussion . . . . .	126
References . . . . .	134
<b>6 Discussion and summary</b>	<b>137</b>
Summary . . . . .	138
Limits of the presented research . . . . .	140
What have we learned at the mechanistic level? . . . . .	142
Future perspectives . . . . .	143
References . . . . .	144

*Contents*

<b>7 Samenvatting</b>	<b>147</b>
-----------------------	------------

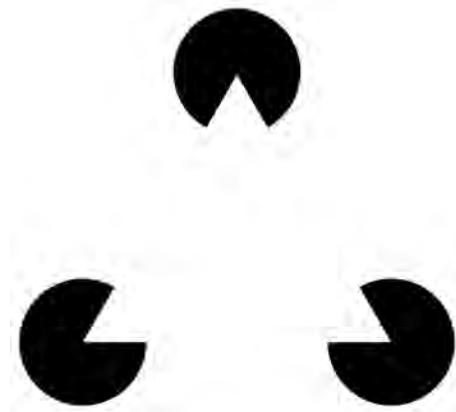
# **1 Introduction**

## *1 Introduction*

The light that falls on our retina is the starting point of the reconstructive process called vision. Photons are emitted or reflected by our surroundings, which may fall on photoreceptors in our eyes, where they are transduced via a chemical cascade into an electrical signal that is transmitted to many subcortical and cortical sites in our brain (Palmer, 2002). Our notion of visual reality depends entirely on the electrical stimulation that is received by these sites from the photoreceptors in our eyes. We are unaware of the interpretative, reconstructive process that underlies the ‘image’ we experience; which is based on neural computations that have been constrained by the phylogeny of the humans species (Nathans et al., 1986a; Nathans et al., 1986b; Nathans, 1999), its ontogeny (Yonas et al., 1993; Sen et al., 2001), and by experience accumulated by the individual (Thorndike, 1901; Thorndike and Woodworth, 1901; Gibson and Gibson, 1955; Shiffrin and Schneider, 1977). These constraints profoundly determine the way in which sensory inputs are processed, and reflect adaptations of the species to suit environmental challenges. For example, bees are unable to see red colors like humans do, however the former’s visual spectrum extends into the ultraviolet range which gives them an advantage in recognizing flowers (Menzel and Blakers, 1975). The present thesis will focus on a class of constraints in the visual system that allows us to perceive surfaces and their boundaries. Surfaces can be defined by multiple cues, including brightness, color, texture and depth (Palmer, 2002). However, because of the cluttered nature of most visual scenes, surfaces and their boundaries are often occluded by surfaces nearer to the observer. The visual system hence faces the task of reconstructing a coherent and behaviorally relevant representation of the surfaces based on partial information. In this reconstructive process, the system selects from an infinite pool of solutions the one that is most likely given constraints about the physical world represented in our brain. The constraints implemented in the visual system compute an automatic assessment of the probability of a particular layout of visual surfaces in the physical world given a partial sensory input on the retina (Lotto et al., 1999; Purves et al., 2004; Yang and Purves, 2004; Purves and Howe, 2005).

## **Visual illusions as a tool to study mechanisms of surface perception**

Because of the filtering of the physical input by constraint-driven visual processes, the resulting representations deviate from the physical input imping-

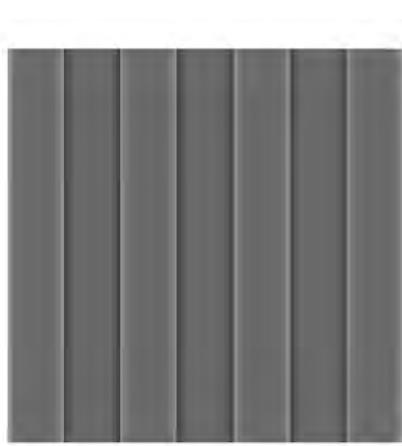


**Figure 1.** Kanizsa triangle.

The alignment of local cues causes the completion of boundaries and the perception of an illusory triangular surface

ing on the retina. The way in which the visual system uses visual constraints to reconstruct representations most likely to conform to physical reality can be studied in stimuli generating visual illusions. The Kanizsa triangle is a well-known example of an illusion that is informative for the way in which the visual system uses local cues to reconstruct surfaces (Schumann, 1900; Kanizsa, 1955; Kanizsa and Gerbino, 1982). Observers typically report a completion of contours along the aligned segments of the inducers. These completed contours form a closed polygon which is perceived as a surface by an observer (Figure 1). Based on sparse information (alignment of local cues) a representation of a visual scene is reconstructed generating the impression of a surface in front of a set of objects in the background. Specifically, observers report surface boundaries that are physically absent, and an impression of enhanced brightness within the confines of the induced surface. Hence, the illusion reveals that in light of the available cues, one of two possible interpretations is highly favored. The display can be seen as containing 4 surfaces (one in the foreground occluding 3 other background surfaces) or as containing 3 surfaces (in which local cues are aligned by some unknown factor). Quite clearly, the visual system favors the former interpretation, despite the physical absence of a surface in the foreground. This indicates that the visual system is heavily biased to interpret the alignment of local cues in the display as evidence for the presence of a surface in the foreground. This visual illusion is thought to reveal the constraints that guide the daily life perception of surfaces in visual scenes.

## *1 Introduction*



**Figure 2.** Craik-O'Brien-Cornsweet (COC) illusion

All vertical bands are of equal luminance but appear perceptually dissimilar due to the luminance ramps near the edges of the bands. Bands with a black luminance ramp near their edge appear darker, bands with a white luminance ramp near their edge appear lighter.

In the Kanizsa triangle, the illusory borders appear to ‘contain’ increased brightness within the induced surface. This suggests important interactions between surface and boundary representations. A visual illusion that explicitly shows the importance of borders in the reconstruction of surface information is the Craik O’Brien Cornsweet (COC) illusion (O’Brien, 1959; Craik, 1966; Cornsweet, 1970; Lotto et al., 1999; Purves et al., 1999; Purves et al., 2004; Purves and Howe, 2005). In the COC illusion, abutting surfaces of the same luminance are perceived as different in brightness (Figure 2). In this stimulus, jigsaw luminance ramps along vertical boundaries create neighboring patches that are physically identical, except that one set is bounded by white edges, and the other by dark ones. The patches bounded by white edges are perceived as brighter than patches bounded by dark edges, in absence of physical differences. Hence, local luminance near the boundaries of the vertical patches has a profound impact on perceived surface brightness.

The two examples discussed reveal strategies the visual system uses to reconstruct surfaces and boundaries from fragmentary retinal input. These illusions have been used frequently as a tool to study the neural correlates of processes underlying surface and boundary perception. A number of landmark neurophysiological studies have revealed neuronal correlates of illusory boundaries and surfaces. Below, these studies will be placed in the

context of other research that has revealed functional subdivisions in the early visual system relevant for surface and boundary processing.

## **Boundary and surface representations: Neurophysiology**

A long tradition of neurophysiological and anatomical research (for review see Sincich and Horton, 2005b, a) has identified two processing streams that differ functionally in a large number of aspects, such that one stream may contribute more to boundary processing, and the other more to surface processing. In V1, interblobs contain neurons that are often orientation selective and that are better stimulated by luminance gradients than by homogenous stimuli, or by isoluminant color gradients. Blobs have been reported to contain less oriented neurons, are often tuned to (very) low spatial frequencies (SF), and (therefore) are well stimulated by homogenous surfaces across their receptive fields (RF). In V2, pale stripes, which receive input from V1 interblobs, contain neurons that predominantly show orientation selectivity, and that produce responses to illusory boundaries (von der Heydt et al., 1984; von der Heydt and Peterhans, 1989; Heitger et al., 1992; Heitger et al., 1998). Thin stripes in V2, which receive input from V1 blobs, contain neurons that show less orientation selectivity and that are sensitive to brightness of homogenous surfaces. For example, Roe et al. (Hung et al., 2001; Roe et al., 2005) have shown modulations of activity in thin stripe neurons in accordance with the illusory brightness changes in the COC stimulus. These data suggest that there are two separate streams in V1 and V2, each of which may have specific contributions to boundary and surface perception. The anatomical data also reveal elaborate lateral connectivity that would permit interpolation of information within subsystems, as well as interactions between them. The neurophysiological data reviewed here also suggest that neuronal activity in V1 is more closely linked to physical properties of local cues, while neuronal activity in V2 may be more closely linked with subjective perception of contours and surfaces.

Note that the responses of neurons that seem to contribute to the perception of visual illusions are driven by local elements outside the confines of the classical RF. The classical RF is traditionally defined as the region in the visual field where a stimulus has to be placed to activate the recorded cell (e.g., Sherrington, 1906; Hartline, 1938; Kuffler, 1953; Hubel and Wiesel, 1962). The finding that neural responses that correlate with the perception of visual illusions are generated from outside the classical RF is in line with

## *1 Introduction*

a long tradition of neurophysiological research showing modulatory effects from extra-RF stimuli (placed in the so-called ‘surround’). For example, several studies have shown in cat area 17 that the response to a small line segment inside the classical RF can be modulated by aligned line segments in the RF surround (Maffei and Fiorentini, 1976; Westheimer et al., 1976; Westheimer and McKee, 1977; Nelson and Frost, 1978, 1985; Gilbert et al., 1990; for review: Lamme and Spekreijse, 1999). The surround elements when presented alone did not result in a measurable response. The anatomical substrate of the surround influences may involve both lateral connections and feedback (Gilbert, 1977; Sirosh et al., 1996; Angelucci and Bressloff, 2006). Hence, the surround of the classical RF is important for contextual interactions among local stimuli, and the responses associated with the perception of illusory boundaries and surfaces can be considered as resulting from a specific kind of contextual interactions.

## **Computational Models**

Three main classes of computational models can be distinguished, each of which have their own assumptions on the contribution of different anatomically and functionally defined subsystems contributing to surface and boundary perception.

According to interpolation theory, surface reconstruction involves active interpolation of surface features from a surface’s edges inward. In this theory, the neural substrate of surface-related spreading exists in early visual areas that are retinotopically organized. Boundary representations would initiate and contain the spreading of surface information. This implies a two-stage process. First, contour interpolation mechanisms use local discontinuities to construct boundary representations. Second, surface interpolation mechanisms use measurements close to contours as seed points for the spread of surface feature, while inhibitory signals emanating from boundary representations contain spreading activity within a surface’s cortical projection. The basic ideas of these models have been formulated more than half a century ago (Fry, 1948; Walls, 1954; Gerrits et al., 1966; Gerrits and Vendrik, 1970), and have formed the basis for a productive line of more contemporary research (Davidson and Whiteside, 1971; Grimson, 1982; Grossberg, 1987b, a; Todorovic, 1987; Grossberg and Todorovic, 1988; Arrington, 1994; Gove et al., 1995; Pessoa et al., 1995; Grossberg, 1997; Grossberg and Raizada, 2000; Neumann et al., 2001; Grossberg, 2003; Grossberg and Seitz, 2003).

Second, spatial frequency filtering theory proposes that surface representations can be derived directly from low spatial frequency information (McCourt, 1982; Stroemer et al., 1984; Purves et al., 1999; Dakin and Bex, 2003; Blakeslee et al., 2005; Blakeslee and McCourt, 2008). This theory is compatible with a correlate of surface perception in retinotopic visual areas. However, low spatial frequencies are processed before higher ones, and therefore, surface representations might emerge prior to the completion of precise boundary representations, which require processing of high spatial frequencies (Hughes et al., 1996).

Third, Dennett (1991) proposed a symbolic encoding theory in which surface qualities are not encoded at all in retinotopic areas. Instead, they are encoded implicitly (symbolically) by a lack of discontinuities between identified surface boundaries. For example, if at the inside of a surface's contour the luminance is high, and at the outside it is low, the entire area within the surface's boundaries is 'labeled' as bright, as long as there is no evidence for other discontinuities inside the surface's representation. Hence, no activity is expected in lower visual areas corresponding to the perceived aspects of surfaces. In this view, surface encoding may take place entirely at a higher, non-retinotopic level in the visual system where surface representations might be an integral aspect of object encoding (Desimone and Ungerleider, 1989; Wang et al., 1996; Biederman, 2000; Haxby et al., 2001; Kayaert et al., 2005).

## **Research question and paradigms**

The goal of the research presented in this thesis was to use fMRI in humans to determine whether there is a signal related to the perception of surfaces. Specifically, we aimed to investigate whether there is a surface-related signal within the confines of its retinotopic projection within early visual areas. Investigating this question is important, because by answering it we will provide support for either the interpolation (and filtering) models, or symbolic theories. Note that the terms 'filling-in' and 'completion' are used to refer to the perceptual effects, and interpolation to the underlying (neural) mechanism.

Although surface perception is linked with boundary perception, this thesis will focus on mechanisms of surface perception (the perceptual aspect in between boundaries). A brightness induction paradigm will be used to study mechanisms of surface brightness perception, and the Troxler paradigm to study mechanisms of surface texture perception. Brightness was chosen as

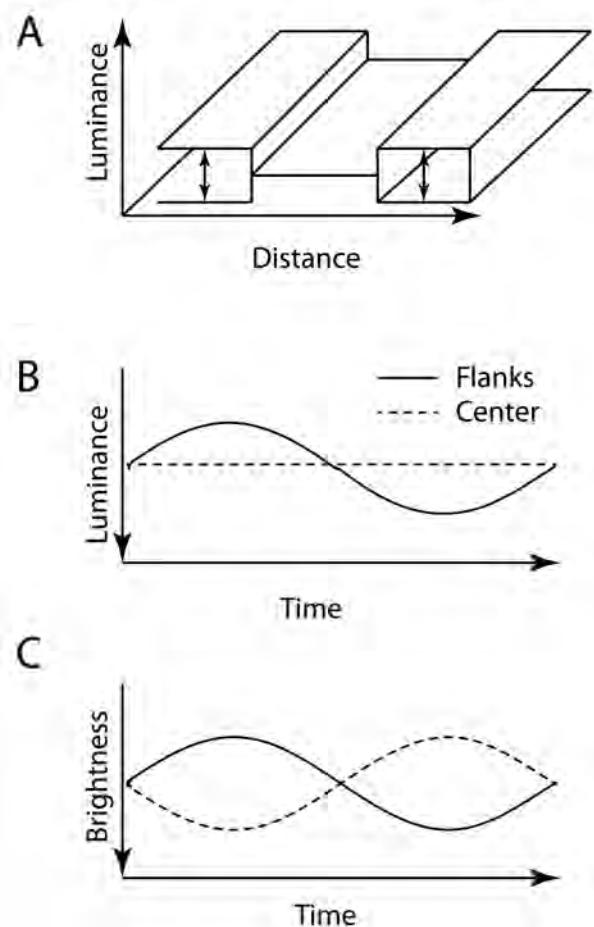
## *1 Introduction*

a topic of study, because it is a well-studied aspect of surfaces and because the contribution of early visual cortex to the perception of surface brightness is currently heavily debated. Texture was chosen because it so far has received less attention and because the few studies that have been published show contradictory results.

### **Surface brightness perception studied via an induction paradigm**

The presented research will focus on a dynamic version of the brightness induction paradigm (Figure 3). In this paradigm (Paradiso and Nakayama, 1991; Rossi et al., 1996; Rossi and Paradiso, 1999), a central grey region of constant luminance is flanked by two regions where luminance is modulated dynamically. The central region is referred to as the probing region, as in that region the presence or absence of an illusion is 'probed'. The flanking regions will be referred to as the inducers, as they 'induce' the illusion. Observers consistently report a modulation of brightness in counterphase to the luminance modulation in inducers.

There are two basic neurophysiological mechanisms that may account for this illusion. One possibility is that neurons with RFs that are sensitive to luminance gradients across their RFs (e.g., simple cells) generate a signal that propagates in the representation of the probing region. The counterphase aspect of the induced brightness may be explained by simple cells with RFs positioned along the border between probing region and inducer region with their OFF subregion inside the inducer region and their ON subregion inside the physically constant probing region. This basic circuitry is in agreement with interpolation models (Davidson and Whiteside, 1971; Grimson, 1982; Grossberg, 1987b, a; Todorovic, 1987; Grossberg and Todorovic, 1988; Arrington, 1994; Gove et al., 1995; Pessoa et al., 1995; Grossberg, 1997; Grossberg and Raizada, 2000; Neumann et al., 2001; Grossberg, 2003; Grossberg and Seitz, 2003), which emphasize that the subjective aspect of surfaces is driven by spread of information from local boundary across the surface. A second possibility is that there is a mutual inhibition between neurons that are sensitive to homogenous regions of luminance-defined surfaces. This basic idea could be seen as compatible with low-spatial frequency analysis proposed in spatial filtering theory (McCourt, 1982; Stromeier et al., 1984; Purves et al., 1999; Dakin and Bex, 2003; Blakeslee et al., 2005; Blakeslee and McCourt, 2008). Finally, it is possible that the brightness of surfaces is computed in high-level areas



**Figure 3.** Brightness induction stimulus modified from Rossi and Paradiso, (1999)  
A) A central grey region of constant luminance (probing region) is flanked by two regions where luminance is modulated dynamically from white to black (inducer regions).  
B) The physical luminance timecourse of the flanker regions (solid line) and central region (dash line).  
C) Perceived brightness of the flanker regions (solid line) and central region (dash line).

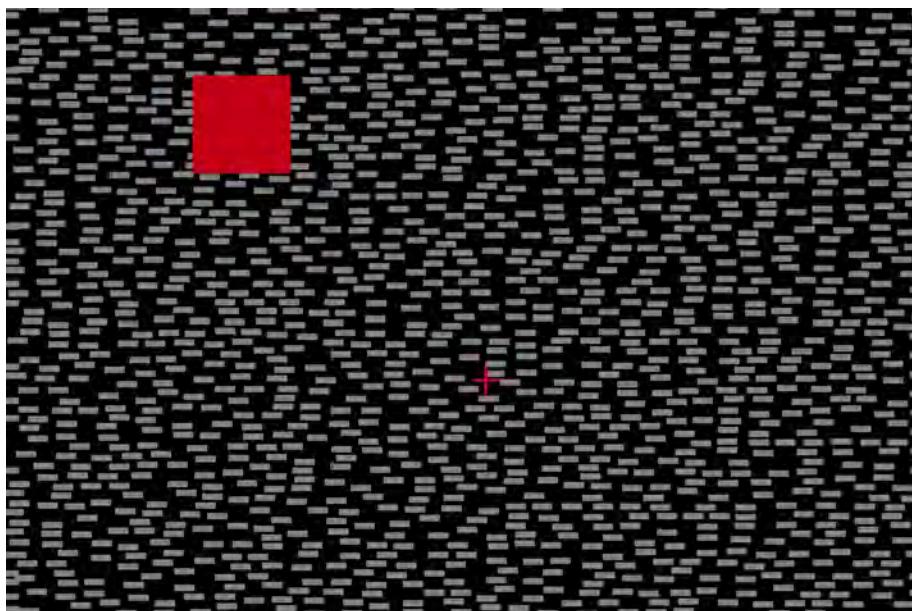
## *1 Introduction*

without the need of an explicit representation in retinotopic areas (Dennett, 1991).

Current neurophysiological research has been limited to area V1, where a very small neural signal has been found correlated with dynamically induced brightness (Rossi et al., 1996; Rossi and Paradiso, 1999). However, fMRI studies have produced conflicting evidence. Cornelissen et al. (2006) reported an absence of any correlates of brightness induction, while Perezvera et al. (2008) did find a correlate with a very similar paradigm. Hence, the fMRI study of brightness induction presented in this thesis was driven by the absence of information on potential extrastriate contributions to surface brightness perception and by the general controversy of the contribution by early visual cortex to surface brightness perception. We also aimed to collect data that could inform computational models of surface brightness perception.

### **Texture surface perception studied via a Troxler paradigm**

In 1804, a set of pioneering experiments on perceptual fading was conducted by Ignaz Paul Vital Troxler and published in an article in Ophthalmologische Bibliothek titled “Über das Verschwinden gegebener Gegenstände innerhalb unseres Gesichtskreises” (On the disappearance of given objects from our visual field) (Troxler, 1804; Iselin, 2005). In this article, Troxler reported the observation that by fixating one's gaze on a fixation point a peripherally presented figure disappears from awareness after a few seconds (Troxler fading). This phenomenon can be experienced in Figure 4, by steadily fixating the red fixation cross, which will lead to a fading of the red square and its replacement by the surrounding texture (perceptual filling-in). This type of delayed filling-in of a figure by its background has been observed in the domains of brightness, color, static and dynamic texture, and motion. The delay before filling-in takes place is shorter the more efficient image stabilization is on the retina. Hence, observers that are capable of fixating better will experience perceptual filling-in more convincingly (Martinez-Conde et al., 2006). Very fast filling-in (as quickly as 80 ms) can be achieved when image stabilization is (near-perfect), which has been accomplished by strategies ranging from contact-lens-cameras, inducing an afterimage on the retina, and monitoring eye movements and using them to compensate for fixation errors (Riggs and Ratliff, 1952; Ditchburn and Ginsborg, 1953; Riggs et al., 1953; Gerrits et al., 1966; Yarbus, 1967; Coppola and Purves, 1996). In the present thesis, the stimulus used to elicit filling-in will consist



**Figure 4.** Texture Filling-in stimulus

By steadily fixating the central cross, the peripheral red square will disappear from awareness and will become filled-in by the line texture after a few seconds.

of a grey figure equiluminant with a surrounding dynamic line texture, and stabilization will be achieved by prolonged fixation in natural conditions.

Because of the analogy of the experience of filling-in across a figure in the Troxler paradigm with filling-in across the blindspot (Mariotte, 1747; Andreae, 1970; Ramachandran, 1992b, a; Grzybowski and Aydin, 2007), Ramachandran suggested that during fixation the figure played the role of an artificially induced scotoma (Ramachandran and Gregory, 1991; Spillmann and Kurtenbach, 1992; Ramachandran et al., 1993). However, considering these two examples of filling-in in the context of the interpolation model shows that this analogy is only partially correct. Interpolation theory suggests that in normal vision, as soon as a saccade lands on a point, boundaries are reconstructed which contain inward spread of surface feature initiated at the boundaries. Surface feature spread is thought to be contained within appropriate retinotopic boundaries by an inhibitory signal generated by the boundary representations (and possibly other sources related to figure-ground segregation) (De Weerd, 2006). Because of the fast rate of successive fixations during normal vision, boundary representations are highly active during the processing of rapidly occurring, successive

## *1 Introduction*

retinal images. It is notable that in the domain of brightness there is compelling psychophysical (Paradiso and Nakayama, 1991; Davey et al., 1998) and neurophysiological evidence (De Weerd et al., 1995) for spread initiated and contained by the boundary. However, in the artificial condition of maintained fixation, boundary representations are thought to adapt, leading to a decline of inhibition, and a spread of surface features outside retinotopically appropriate boundaries. This would lead to the percept of filling-in in the Troxler paradigm, and also explains the delayed nature of the experience. By contrast, filling-in across the blind spot is immediate as there are no boundary representations that have to adapt to permit the spreading of surface feature into the cortical region occupied by the natural blind spot. The main alternative to the interpolation model is the symbolic model, according to which the fading of the figure's boundaries would be the sole reason for the subjective disappearance of the figure from view. In this model, due to the adaptation of the figure's boundaries, high-level surface representations are updated to ignore the area physically occupied by the figure, thereby automatically representing the background surface without any discontinuities.

Little is known about the neural correlates of surface texture filling-in. It was first investigated in monkeys (De Weerd et al., 1995), by means of single cell recording. In this study, De Weerd et al. (1995) found strong activity increases in extrastriate neurons of V2 and V3 with RFs over a grey region in a dynamic texture during maintained fixation away from the figure. These activity increases coincided with the timing of filling-in in human observers exposed to the same stimulus conditions. Surprisingly, in recent MEG and fMRI studies, the claim has been made that in human visual cortex, filling-in by dynamic texture is associated with activity decreases. In light of these contradictory conclusions drawn from the two sets of studies, we chose to reinvestigate the neural correlates of texture filling-in using stimulus conditions closely resembling the ones used in the monkey recording studies. In a larger perspective, we aimed to collect data that could inform computational models of perception.

## **Methodological considerations**

The aim of this thesis is to identify a neural signal related to surface perception using 3T fMRI in human participants, in two different paradigms. The neurophysiological evidence suggests that the signal elicited in the early visual system by physical homogenous surfaces is small (Murphy et al., 2007),

which makes it likely that an illusion that might reveal neural interpolation processes will correspond to an even smaller signal. In addition, some of the evidence suggests that this neural interpolation signal, if it exists, may be stronger in extrastriate than striate cortex. Unfortunately, in extrastriate cortex, cortical magnification is lower, and RF sizes are larger (Kornhuber and Deecke, 1965) than in striate cortex. In combination with limits in spatial resolution of 3T fMRI, the sensitivity to pick up the signal of interest therefore will be comparatively smaller precisely in those brain regions where it is most likely to occur.

Therefore, one of the challenges for the research planned in the present thesis was to tailor the data acquisition and analysis to increase the likelihood to pick up the signal of interest. A straightforward way to increase this likelihood is by increasing the number of trials a participant has to undergo. By averaging more trials, the signal to noise ratio increases assuming that the noise in the signal is Gaussian, not temporarily correlated and varies randomly across trials. In addition to merely increase power by increasing numbers of trials, we also designed procedures for informed temporal alignment of relevant stimulus periods prior to averaging, and precise procedures for functional spatial alignment prior to averaging.

The approach used for temporal averaging is similar to the approach commonly used to find small amplitude components in EEG waveforms, commonly referred to as event-related averaging. Typically, event-related averaging involves that signals are time-locked to the stimulus onset, and then averaged. This averaging reduces the noise and unveils the components present underneath the large amplitude waves in the signal. In the experiments investigating texture filling-in, the 'stimulus' is the onset of an illusion defined by a specific criterion as measured while participants are being scanned. For spatial alignment prior to averaging, a new method was devised inspired by techniques that are currently used to align spectrograms yielded by molecular imaging experiments. The next section briefly introduces the chapters that describe the research triggered by the scientific questions that have been outlined here.

## **Thesis overview**

Chapter 2 offers a report on the development and testing of an efficient method for spatial alignment prior to fMRI data averaging. This research is presented first, as tackling the problem of alignment is a precondition for answering the empirical questions formulated in this Introduction. Chapter 3

## *1 Introduction*

presents the results from an experiment that tested the existence of an fMRI signal associated with surface brightness perception in early visual cortex in a brightness induction paradigm. Chapter 4 shows a simple forward neural network model that successfully generated a dataset that could be compared with the empirical data from Chapter 3. This comparison was used to evaluate the truth value of the assumptions underlying the model. Chapter 5 contains the results from an fMRI study testing the existence of correlates of surface texture perception in a Troxler paradigm. Finally, Chapter 6 contains a summary of the findings and a discussion of the implications of the experiments for the models put forward in the Introduction. The chapter will end with an evaluation of the strengths and weaknesses of the type of data reported in this thesis, and an outline of perspectives for further research.

## **References**

- Andreae, H. (1970). Edme Mariotte-born 350 years ago. *Quintessenz* 21, 153-155.
- Angelucci, A., and Bressloff, P.C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Prog Brain Res* 154, 93-120.
- Arrington, K.F. (1994). The temporal dynamics of brightness filling-in. *Vision Res* 34, 3371-3387.
- Biederman, I. (2000). Recognizing depth-rotated objects: A review of recent research and theory. *Spatial Vision* 13, 241-253.
- Blakeslee, B., and McCourt, M.E. (2008). Nearly instantaneous brightness induction. *Journal of Vision* 8, 1-8.
- Blakeslee, B., Pasieka, W., and McCourt, M.E. (2005). Oriented multiscale spatial filtering and contrast normalization: A parsimonious model for brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Res* 45, 607-615.
- Coppola, D., and Purves, D. (1996). The extraordinarily rapid disappearance of entoptic images. *Proc Natl Acad Sci U S A* 93, 8001-8004.
- Cornelissen, F.W., Wade, A.R., Vladusich, T., Dougherty, R.F., and Wandell, B.A. (2006). No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex. *J Neurosci* 26, 3634-3641.
- Cornsweet, T.N. (1970). *Visual Perception* (New York: Academic).
- Craik, K.J.W. (1966). *The Nature of Psychology* (Cambridge: Cambridge University Press).
- Dakin, S.C., and Bex, P.J. (2003). Natural image statistics mediate brightness 'filling

## References

- in'. Proc Biol Sci 270, 2341-2348.
- Davey, M.P., Maddess, T., and Srinivasan, M.V. (1998). The spatiotemporal properties of the Craik-O'Brien-Cornsweet effect are consistent with 'filling-in'. Vision Res 38, 2037-2046.
- Davidson, M., and Whiteside, J.A. (1971). Human brightness perception near sharp contours. J Opt Soc Am 61, 530-536.
- Dennett, D.C. (1991). Consciousness explained (Boston: Little, Brown and Company).
- von der Heydt, R., and Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. J Neurosci 9, 1731-1748.
- von der Heydt, R., Peterhans, E., and Baumgartner, G. (1984). Illusory contours and cortical neuron responses. Science 224, 1260-1262.
- Desimone, R., and Ungerleider, L.G. (1989). Neural mechanisms of visual processing in monkeys. In Handbook of Neuropsychology, F. Boller, and J. Grafman, eds. (Amsterdam: Elsevier, Science Publishers B.V. (Biomedical Division)), pp. 267-299.
- Ditchburn, R.W., and Ginsborg, B.L. (1953). Involuntary eye movements during fixation. J Physiol 119, 1-17.
- Fry, G.A. (1948). Mechanisms subserving simultaneous brightness contrast. Am J Optom Arch Am Acad Optom 25, 162-178.
- Gerrits, H.J., De Haan, B., and Vendrik, A.J. (1966). Experiments with retinal stabilized images. Relations between the observations and neural data. Vision Res 6, 427-440.
- Gerrits, H.J., and Vendrik, A.J. (1970). Simultaneous contrast, filling-in process and information processing in man's visual system. Exp Brain Res 11, 411-430.
- Gibson, J.J., and Gibson, E.J. (1955). Perceptual learning; differentiation or enrichment? Psychol Rev 62, 32-41.
- Gilbert, C.D. (1977). Laminar differences in receptive field properties of cells in cat primary visual cortex. J Physiol 268, 391-421.
- Gilbert, C.D., Hirsch, J.A., and Wiesel, T.N. (1990). Lateral interactions in visual cortex. Cold Spring Harb Symp Quant Biol 55, 663-677.
- Gove, A., Grossberg, S., and Mignolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. Visual Neuroscience 12, 1027-1052.
- Grimson, W.E. (1982). A computational theory of visual surface interpolation. Philosophical Transactions of the Royal Society B: Biological Sciences 298, 395-427.
- Grossberg, S. (1987a). Cortical dynamics of three-dimensional form, color, and brightness perception: I. Monocular theory. Percept Psychophys 41, 87-116.
- Grossberg, S. (1987b). Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory. Percept Psychophys 41, 117-158.
- Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. Psychol Rev 104, 618-658.
- Grossberg, S. (2003). Filling-in the forms: Surface and boundary interaction in visual

## *1 Introduction*

- cortex. In Filling-in: From Perceptual Completion to Skill LEarning, L. Pessoa, and P. De Weerd, eds. (Oxford: University Press), pp. 13-37.
- Grossberg, S., and Raizada, R.D. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Res* 40.
- Grossberg, S., and Seitz, A. (2003). Laminar development of receptive fields, maps and columns in visual cortex: the coordinating role of the subplate. *Cereb Cortex* 13, 852-863.
- Grossberg, S., and Todorovic, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: a unified model of classical and recent phenomena. *Percept Psychophys* 43, 241-277.
- Grzybowski, A., and Aydin, P. (2007). Edme Mariotte (1620-1684): Pioneer of Neurophysiology. *Surv Ophthalmol* 52, 443-451.
- Hartline, H.K. (1938). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *Am J Physiol* 121, 400-415.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., and Peietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293.
- Heitger, F., Von der Heydt, R., Peterhans, A., Rosenthaler, L., and Kübler, O. (1998). Simulation of neural contour mechanisms: representing anomalous contours. *Image and Vision Computing* 16, 409-423.
- Heitger, F., Rosenthaler, L., von der Heydt, R., Peterhans, E., and Kubler, O. (1992). Simulation of neural contour mechanisms: from simple to end-stopped cells. *Vision Res* 32, 963-981.
- Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160.
- Hughes, H.C., Nozawa, G., and Kitterle, F. (1996). Global Precedence, Spatial Frequency Channels, and the Statistics of Natural Images. *Journal of Cognitive Neuroscience* 8, 197-230.
- Hung, C.P., Ramsden, B.M., Chen, L.M., and Roe, A.W. (2001). Building surfaces from borders in Areas 17 and 18 of the cat. *Vision Res* 41, 1389-1407.
- Iselin, H.U. (2005). Identifying Troxler. *British Journal of Ophthalmology* 90, 250.
- Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia* 49, 7-30.
- Kanizsa, G., and Gerbino, W. (1982). Amodal completion: Seeing or thinking? In Organization and Representation in Perception, J. Beck, ed. (Hillsdale NJ: Lawrence Earlbaum Associates), pp. 167-190.
- Kayaert, G., Biederman, I., Op de Beeck, H.P., and Vogels, R. (2005). Tuning for shape dimensions in macaque inferior temporal cortex. *European Journal of Neuroscience* 22, 212-224.

## References

- Kornhuber, H.H., and Deecke, L. (1965). Changes in the Brain Potential in Voluntary Movements and Passive Movements in Man: Readiness Potential and Reafferent Potentials. *Pflugers Arch Gesamte Physiol Menschen Tiere* 284, 1-17.
- Kuffler, S.W. (1953). Discharge patterns and functional organization of mammalian retina. *J Neurophysiol* 16, 37-68.
- Lamme, V.A., and Spekreijse, H. (1999). Contextual Modulation in Primary Visual Cortex and Scene Perception. In *The New Cognitive Neurosciences*, M.S. Gazzaniga, ed. (Cambridge, MA: The MIT Press).
- Lotto, R.B., Williams, S.M., and Purves, D. (1999). An empirical basis for Mach bands. *Proc Natl Acad Sci U S A* 96, 5239-5244.
- Maffei, L., and Fiorentini, A. (1976). The unresponsive regions of visual cortical receptive fields. *Vision Res* 16, 1131-1139.
- Mariotte, E. (1747). Nouvelle Decouverte touchant la Vue, contenue en plusieurs Lettres. In *Oeuvres de Mr. Mariotte, de l'Academie Royale des Sciences*, M. E., ed. (Leiden: Pierre Vander Aa), pp. 496-498.
- Martinez-Conde, S., Macknik, S.L., Troncoso, X.G., and Dyar, T.A. (2006). Microsaccades counteract visual fading during fixation. *Neuron* 49, 297-305.
- McCourt, M.E. (1982). A spatial frequency dependent grating-induction effect. *Vision Res* 22.
- Menzel, R., and Blakers, M. (1975). Color receptors in the bee eye - Morphology and spectral sensitivity. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural and Behavioral Physiology* 108, 11-13.
- Murphy, K., Bodurka, J., and Bandettini, P.A. (2007). How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *Neuroimage* 34, 565-574.
- Nathans, J. (1999). The evolution and physiology of human color vision: insights from molecular genetic studies of visual pigments. *Neuron* 24, 299-312.
- Nathans, J., Piantanida, T.P., Eddy, R.L., Shows, T.B., and Hogness, D.S. (1986a). Molecular genetics of inherited variation in human color vision. *Science* 232, 203-210.
- Nathans, J., Thomas, D., and Hogness, D.S. (1986b). Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science* 232, 193-202.
- Nelson, J.I., and Frost, B.J. (1978). Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Res* 139, 359-365.
- Nelson, J.I., and Frost, B.J. (1985). Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. *Exp Brain Res* 61, 54-61.
- Neumann, H., Pessoa, L., and Hansen, T. (2001). Visual filling-in for computing perceptual surface properties. *Biol Cybern* 85, 355-369.
- O'Brien, V. (1959). Contrast by contour-enhancement. *American Journal of Psychology* 72.

## *1 Introduction*

- Palmer, S.E. (2002). *Vision Science: Photons to Phenomenology*. (Cambridge Mass., MIT Press).
- Paradiso, M.A., and Nakayama, K. (1991). Brightness perception and filling-in. *Vision Res* 31, 1221-1236.
- Pereverzeva, M., and Murray, S.O. (2008). Neural activity in human V1 correlates with dynamic lightness induction. *J Vis* 8, 8 1-10.
- Pessoa, L., Mingolla, E., and Neumann, H. (1995). A contrast- and luminance-driven multiscale network model of brightness perception. *Vision Res* 35, 2201-2223.
- Purves, D., and Howe, C.Q. (2005). *Perceiving Geometry: Geometrical Illusions Explained by Natural Scene Statistics* (New York: Springer).
- Purves, D., Shimpi, A., and Lotto, R.B. (1999). An empirical explanation of the cornsweet effect. *J Neurosci* 19, 8542-8551.
- Purves, D., Williams, S.M., Nundy, S., and Lotto, R.B. (2004). Perceiving the intensity of light. *Psychol Rev* 111, 142-158.
- Ramachandran, V.S. (1992a). Blind spots. *Sci Am* 266, 86-91. Ramachandran, V.S. (1992b). Filling in the blind spot. *Nature* 356, 115.
- Ramachandran, V.S., and Gregory, R.L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature* 350, 699-702.
- Ramachandran, V.S., Gregory, R.L., and Aiken, W. (1993). Perceptual fading of visual texture borders. *Vision Res* 33, 717-721.
- Riggs, L., and Ratliff, F. (1952). The effects of counteracting the normal movements of the eye. *J Opt Soc Am* 42, 872-873.
- Riggs, L.A., Ratliff, F., Cornsweet, J.C., and Cornsweet, T.N. (1953). The disappearance of steadily fixated visual test objects. *J Opt Soc Am* 43, 495-501.
- Roe, A.W., Lu, H.D., and Hung, C.P. (2005). Cortical processing of a brightness illusion. *Proc Natl Acad Sci U S A* 102, 3869-3874.
- Rossi, A.F., and Paradiso, M.A. (1999). Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *J Neurosci* 19, 6145-6156.
- Rossi, A.F., Rittenhouse, C.D., and Paradiso, M.A. (1996). The representation of brightness in primary visual cortex. *Science* 273, 1104-1107.
- Saad, Z.S., Ropella, K.M., DeYoe, E.A., and Bandettini, P.A. (2003). The spatial extent of the BOLD response. *Neuroimage* 19, 132-144.
- Schumann, F. (1900). Beiträge zur Analyse der Gesichtswahrnehmungen. Erste Abhandlung. Einige Beobachtungen über die Zusammenfassung von Gesichtseindrücken zu Einheiten. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* 23, 1-32.
- Sen, M.G., Yonas, A., and Knill, D.C. (2001). Development of infants' sensitivity to surface contour information for spatial layout. *Perception* 30, 167-176.
- Sherrington, C.S. (1906). *The integrative action of the nervous system* (New York: C Scribner and Sons).

## References

- Shiffrin, R.M., and Schneider, W. (1977). Controlled and automatic human information processing: II. perceptual learning, automatic attending, and a general theory. *Psychol Rev* 84, 127-190.
- Sincich, L.C., and Horton, J.C. (2005a). The circuitry of V1 and V2: integration of color, form, and motion. *Annu Rev Neurosci* 28, 303-326.
- Sincich, L.C., and Horton, J.C. (2005b). Input to V2 thin stripes arises from V1 cytochrome oxidase patches. *J Neurosci* 25, 10087-10093.
- Sirosh, J., Miikulainen, R., and Choe, Y. (1996). Lateral Interactions in the Cortex: Structure and Function.
- Spillmann, L., and Kurtenbach, A. (1992). Dynamic noise backgrounds facilitate target fading. *Vision Res* 32, 1941-1946.
- Stromeyer, C.F., Kronauer, R.E., Madsen, J.C., and Klein, S.A. (1984). Opponent-Movement mechanisms in human vision. *Journal of the Optical Society of America, Optic and Image Science* 1.
- Thorndike, E.L. (1901). Adaptation in Vision. *Science* 14, 221.
- Thorndike, E.L., and Woodworth, R.S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychol Rev* 8, 247-261.
- Todorovic, D. (1987). The Craik-O'Brien-Cornsweet effect: New varieties and their theoretical implications. *Perception & Psychophysics* 42.
- Troxler, I.P.V. (D) (1804). Über das Verschwinden gegebener Gegenstände innerhalb unseres Gesichtskreises. In *Ophthalmologische Bibliothek*, K. Himly, and J.A. Schmidt, eds. (Jena, Germany: Fromann), pp. 51-53.
- Walls, G.L. (1954). The filling-in process. *Am J Optom Arch Am Acad Optom* 31, 329-341.
- Wang, G., Tanaka, K., and Tanifuji, M. (1996). Optical imaging of functional organization in the monkey inferotemporal cortex. *Science* 272, 1665-1668.
- De Weerd, P. (2006). Perceptual filling-in: More than the eye can see. *Prog Brain Res* 154, 227-245.
- De Weerd, P., Gattass, R., Desimone, R., and Ungerleider, L.G. (1995). Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature* 377, 731-734.
- Westheimer, G., and McKee, S.P. (1977). Spatial configurations for visual hyperacuity. *Vision Res* 17, 941-947.
- Westheimer, G., Shimamura, K., and McKee, S.P. (1976). Interference with line-orientation sensitivity. *J Opt Soc Am* 66, 332-338.
- Yang, Z., and Purves, D. (2004). The statistical structure of natural light patterns determines perceived light intensity. *Proc Natl Acad Sci U S A* 101, 8745-8750.
- Yarbus, A.L. (1967). *Eye movements and vision* (New York: Plenum Press).
- Yonas, A., Knill, D.C., Sen, A., and Bittinger, K.A. (1993). The development in human

## *1 Introduction*

infants of sensitivity to surface contour information for 3-D layout. In *Studies in Perception and Action II*, S.S. Valenti, and J. Pittenger, eds. (Hillsdale, NJ: Lawrence Erlbaum Associates Publishers).

## **2 Calculating group level activity landscapes of spatial fMRI data**

based on:

Jans, B., Peters, J., and De Weerd, P., (2010), Visual spatial attention to multiple locations at once: The jury is still out. *Psychological Review*, 117(2):637-684.

Jans, B., Been, M., Van de Ven, V., Goebel, R., De Weerd, P., Calculating group level activity landscapes of spatial fMRI data. *Submitted*

## **Abstract**

To compare imaging data among groups of participants, the functional data sets must be aligned. Existing alignment methods are based on anatomical criteria, either using sparse landmarks or cortical curvature (cortex-based alignment, CBA) as criteria for alignment. However, because functional maps are only loosely tied to anatomical landmarks, the enhancement of statistical power after alignment is suboptimal, and spatial detail present in individual maps is lost. Here, we propose pattern-based alignment (PBA), a new method that is based on topographical patterns of activity elicited by localizer stimuli. In empirical and simulated datasets, we used signal analysis and other measures to quantify the amount of spatial detail in functional data before and after alignment by CBA and PBA, and found that PBA outperforms CBA. We suggest that the possibility to compute across-participant maps of functional activity showing exquisite spatial detail will make PBA a major new tool in the emerging field of high field fMRI.

## **Introduction**

An important issue in fMRI data analysis is how to aggregate datasets from several subjects (Lange et al., 1999). Averaging over multiple subjects is required to increase statistical power, but often a significant part of this gain in power is lost by poor alignment. This problem is most pronounced in datasets obtained in topographically organized areas, in which fine-grained voxel-wise spatial patterns of data have to be properly aligned. In visual cortex, topography is best documented (Daniel and Whitteridge, 1961; Gattass et al., 1981; Sereno et al., 1995; Engel et al., 1997; Qiu et al., 2006; Dumoulin and Wandell, 2008; Datta and DeYoe, 2009), and the present paper will tackle the issue how to preserve spatial detail in functional data during alignment in early, retinotopic cortex.

A classical anatomical approach to align datasets of multiple subjects is to explicitly register each brain into a standardized space, as done in Talairach (Talairach and Tournoux, 1988) or MNI (Chau and McIntosh, 2005) normalisation. In the Talairach method, the brain is divided in 12 parts based on anatomical landmarks. These parts are then squeezed or stretched to fit in the standard Talairach coordinate system, which is based on a single brain. The MNI (Chau and McIntosh, 2005; Lancaster et al., 2007) normalisation approach is similar to the Talairach approach, except that experimental subjects' brains are aligned to the average of multiple

brains. Thus, in contrast to the Talairach method, in which all brains are aligned to that of a single arbitrarily chosen individual, MNI uses a representative standard. Currently, a well-accepted standard is ICBM152, which has been adopted as a standard by the International Consortium of Brain Mapping. Although MNI has the advantage of aligning individual brains to a more realistic standard compared to the Talairach method, both methods share the limitation that alignment is based on sparse anatomical landmarks.

The approach of alignment based on sparse landmarks is superseded by modern anatomical alignment methods that explicitly align the convex and concave parts of the cortical sheet, gyri and sulci (Fischl et al., 1999; Goebel et al., 2006), also referred to as Cortex Based Alignment (CBA) (Fischl et al., 1999). This is a method that does not rely on a pre-defined template and that permits a much finer degree of anatomical alignment than can be achieved in standard Talairach normalization (Talairach and Tournoux, 1988) or MNI (Chau and McIntosh, 2005; Lancaster et al., 2007). CBA has indeed been shown (Fischl et al., 1999) to substantially improve statistical group results in aggregated functional maps associated with the aligned brains.

The principal limitation of all anatomical alignment methods is that the relationship between function and anatomy across individuals may vary considerably. If this is the case, even the most sophisticated anatomical alignment method will not lead to optimized power after averaging across aligned functional datasets from different subjects. While the dissociation between anatomy and function is well known in high-level centres of the brain (e.g., frontal lobe; Fischl et al., 2008), it plays a role even in early visual retinotopic areas where the relationship between anatomy and function is often believed to be fairly constant. For example, although according to the classical model of Holmes a close correspondence exists between the deepest point of the calcarine fissure and the horizontal meridian representation (Holmes, 1919; Horton and Hoyt, 1991), functional studies have put this correspondence in doubt (Stensaas et al., 1974; Ono et al., 1990; Sereno et al., 1995; Dougherty et al., 2003). In addition, retinotopic representations derived from cytoarchitectonic data suggest a mismatch between anatomy and functional visual areas (Clarke and Miklossy, 1990; Amunts et al., 2000; Wilms et al., 2010). Moreover, the overall size of early retinotopic areas varies greatly among individuals and even between hemispheres of the same individual (Brindley, 1972; Stensaas et al., 1974; Kirson et al., 2008), which

## *2 Activity landscapes fMRI*

limits the effectiveness of anatomical alignment of functional data even in retinotopic visual cortex.

Using anatomical alignment to achieve correspondence of functional areas is a detour that is unnecessary if *a priori* regions of interest (ROIs) can be defined by means of an independent localizer (Brett et al., 2002; Friston et al., 2006; Mitsis et al., 2008). In higher-order visual cortex, specific stimulus sets can be used to identify areas such as MT (Zeki et al., 1991; Watson et al., 1993), the lateral occipital complex (Malach et al., 1995; Grill-Spector et al., 2001), the fusiform face area (Sergent et al., 1992; Kanwisher et al., 1997) or the parahippocampal place area (Aguirre et al., 1998; Ishai et al., 1999). In lower-order retinotopic cortex a localizer stimulus at specific visual field coordinates can be used as an independent localizer for an experimental stimulus in that location. Typically, studies that use a region of interest (ROI) approach average the responses to the experimental stimulus over all voxels in the ROI in a subject, and then average over all subjects, thereby achieving a simple form of functional alignment that bypasses the problems of cortical alignment. A consequence of this approach, however, is that possible spatial patterns in the functional data within ROIs are lost.

The goal of the present paper is to demonstrate a Pattern Based Alignment (PBA) technique that makes it possible to perform a group analysis of fine-grained, spatial (topographical) patterns in functional data. We will make a qualitative and quantitative comparison between anatomically aligned data sets and functionally aligned data sets, using both real and simulated data. The demonstration of PBA is focused on specific spatial patterns of functional data in retinotopically organized visual cortex, but can be expanded to other topographically organized cortex, or to any other experiment in which a specific pattern of functional activation is expected. The method we propose is based on a non-linear morphing of the cortical sheet occupied by a ROI constrained by independent localizers designed to functionally map this ROI. This results in an excellent alignment across subjects of functional data in a normalized anatomical space.

## **Methods**

### **fMRI experiments**

Anatomical T1-weighted (MP-RAGE; TR, 2,250 ms; matrix size, 256×256; 192 slices; in-plane resolution, 1 mm<sup>2</sup>) and functional echo planar imaging (EPI) scans (14 gapless slices, TR/TE, 1,000/30 ms; flip angle 90°; square in plane matrix size of 128<sup>2</sup>; resulting in isotropic 2mm<sup>3</sup> voxels) were ac-

## *Methods*

quired of two subjects on a 3T Siemens Allegra head scanner. Note that the functional slices were positioned oblique, parallel and below the calcarine sulcus, restricting the visual field coverage to the upper quadrants. Functional localizer and retinotopic data were pre-processed in Brainvoyager QX (Brain Innovation, The Netherlands) including slice time correction, 3D correction for head volume movements and temporal filtering, which included linear trend removal and a high-pass filter of 3 cycles per time course. Finally, functional data were resampled and co-registered to the Talairach transformed anatomical data (Goebel et al., 2006). Retinotopic data were acquired according to standard procedures (Sereno et al., 1995; Engel et al., 1997). Polar angle mapping allowed us to identify V1v, V2v, and V3v in the CBA data at the group level as well as in the individual native hemispheres. During the localizer runs 3 eccentricity localizer stimuli (Figure 5A) were presented to the subjects in 3 functional runs per eccentricity, runs lasted 290s and consisted of alternating baseline and localizer blocks with duration of 11s. Experimental stimuli with luminance modulation in the outer ring and inner disk were presented in runs of 460s and contained 32 luminance modulations.

Anatomical data were corrected for field inhomogeneities and volumetrically normalized into Talairach space (Talairach and Tournoux, 1988). A brightness cut-off method was used to identify the white-gray matter boundary, where meshes were modelled to reconstruct the cortical sheet. These operations facilitated the creation of participants' 3D mesh reconstructions. The meshes were subsequently inflated, flattened and corrected for spatial distortions.

### **Cortex based alignment (CBA)**

The CBA procedures consisted of down sampling the number of vertices in the 3d meshes. This resulted in smoothing of the meshes, while preserving their shape. The downsampled meshes were then inflated to spheres and by using gyral and sulcal curvature constraints spheres they were aligned using a multi scale approach. Here in a first alignment step, curvature information is limited to gross curvature features. While the alignment process progresses more detailed curvature information is used and aligned to a moving group average sphere (Goebel et al., 2006). To align both right and left hemispheres, left hemispheres are mirrored.

## Functional localiser analysis and experimental data

Multiple regression analysis approach was used to fit a general linear model (GLM) with box car predictors for each localizer condition (3 conditions; radius inner 3 degrees visual angle  ${}^{\circ}$ ), middle  $9^{\circ}$ , outer ring  $14^{\circ}$ ; see Figure 5A and Figure 5D) to the data. Beta coefficient maps were then created for each localizer condition and used for further alignment analysis. The experimental stimulus had the same dimensions as the localizer stimulus and was built out of two concentric rings around a disk. The outer ring and inner disk were brightness modulated sinusoidal from white to black (duration 1s) followed by a period (of 8s) wherein the rings did not change luminance and modulated back from black to white (duration 1s) while the middle ring was constant black.

## Simulations

Reference data were simulated as a weighted sum of sinoids, the signal  $s$  was then enveloped by a Gaussian.

$$\begin{aligned} t &= 0 \dots 7\pi \\ s &= \sin(t + \pi) + 0.5 \sin(3t + \pi) + 0.5 \sin(9t + \pi) + 0.25 \sin(27t + \pi) \\ s &= s^T * e^{\left(\frac{x-\mu^2}{2\sigma^2}\right)} \end{aligned}$$

## SNR calculation

$$SNR = 10 \log_{10} \left( \frac{\sum_{i=1}^n reference^2}{\sum_{i=1}^n reference - sampled^2} \right)$$

All alignment routines were implemented in Matlab (The Mathworks, Natick, MA, USA) and Brainvoyager QX (Brain Innovation, The Netherlands).

## Results

To illustrate the effectiveness of PBA, we used CBA as a benchmark. Here, we first compare and demonstrate the two algorithms, and then compare the effect of the two algorithms on the alignment of real and simulated functional data.

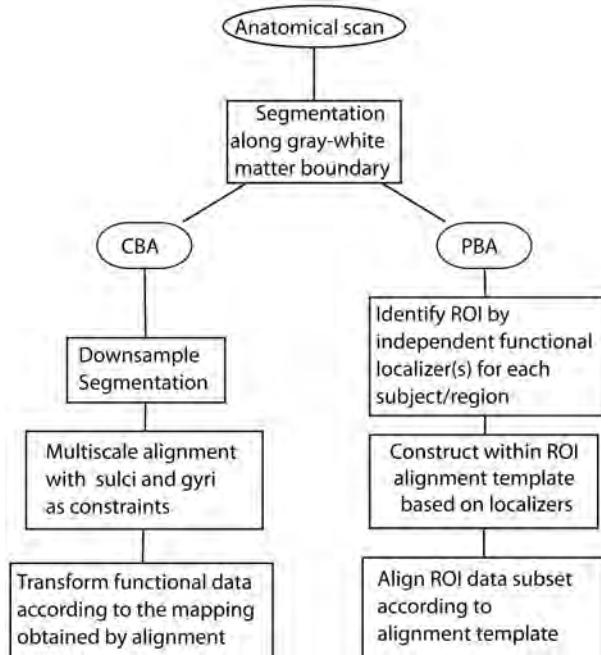
## CBA versus PBA

Both CBA and PBA techniques rely on a surface reconstruction of the cortical sheet on which subsequent alignment steps are done. Nevertheless, already during the early step of surface reconstruction, CBA and PBA differ (Figure 1). For the CBA technique to work a large amount of smoothing is necessary, because spurious individual details can lead to misalignment. Therefore, the reconstructed meshes are downsampled along with the curvature information superimposed on the meshes; this leads to fewer vertices and a less detailed description of the surface curvature. Since the functional data undergo the same transformation as the anatomical data, the fine details in the functional data are lost as well after CBA. By contrast, the goal of PBA is to preserve as much of the detail as possible, and therefore all operations are applied on the native resolution meshes (meshes in which vertices are not decimated).

## PBA algorithm and illustration

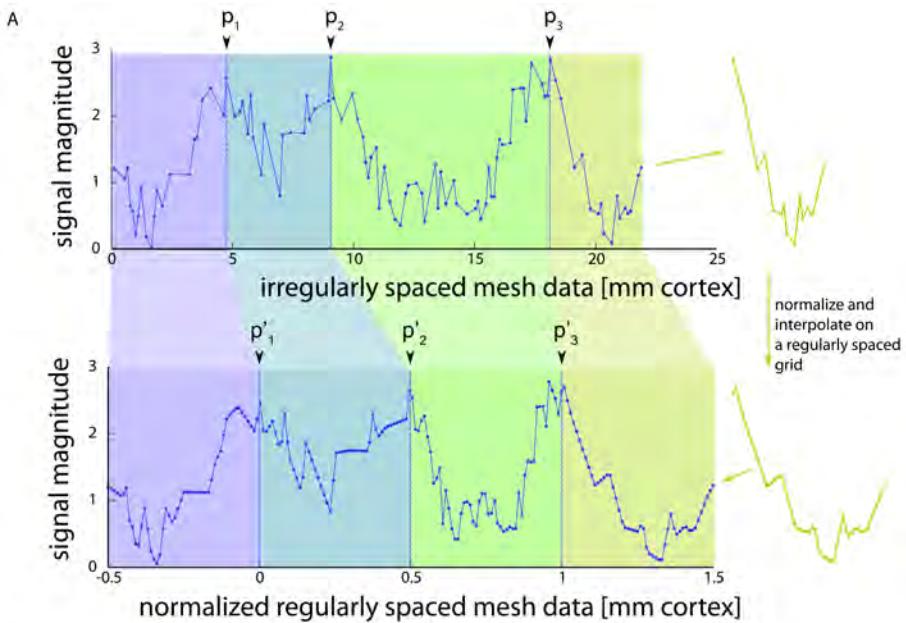
We will first illustrate the effects of different alignment algorithms along a one-dimensional sample of the 2D reconstruction of the cortical sheet of V1, in the left and right hemisphere of a participant. To that aim, we defined a cortical path through the map in each hemisphere, and through an area activated by three localizer annuli (radii of 3°, 9° and 14°). We only considered functional data along this anatomical path for alignment. To combine functional data from ROIs, we collected data from a second set of within-ROI functional localizers. The peaks within the functional data along each anatomical path served as a pattern of landmarks which then were aligned among individual paths, to construct an alignment template. After construction of the template, there is a point-to-point correspondence between the functional data sets along the different cortical paths, and with the template.

The algorithm involved in establishing the correspondence between an arbitrary cortical path and a template is described visually and by pseudo code in Figure 2A-B. The spatial profile of the fMRI signal along the path was divided into segments based on landmarks determined from within-ROI localizers. Each segment was then normalized and interpolated onto a regular grid, and once this was done for all segments, the transformed segments were recombined. As a result of this procedure, signals coming from different anatomical paths from different hemispheres or participants can be compared and averaged (they are functionally aligned). Figure 3



**Figure 1.** Flow chart comparison between CBA and PBA

Flow chart comparing the necessary steps for Cortex Based Alignment (CBA) and Pattern Based Analysis (PBA). After white/grey matter segmentation, process pipelines of CBA and PBA diverge. In CBA, anatomical alignment between brains is obtained by downsampling the surface resulting from segmentation, and by multiscale alignment of gyri and sulci of individual brains. Functional data are transformed accordingly. PBA does not use whole brain information for alignment but rather aligns patterns within ROIs identified by within-ROI functional localizers. Based on functional landmarks within an ROI a template is formed to which individual functional data are aligned.



B

## Pseudocode

```

Define the signal characteristics to align  $p_{1-3}$  (in inset A 3 peaks,  $p_1$ ,  $p_2$ ,  $p_3$ )
number_of_intervals = number_of_peaks + 1;
p = {0,  $p_1$ ,  $p_2$ ,  $p_3$ , end}; % alignment points including start and end

intervallength = 20;

for (interval=1; interval==number_of_intervals; n += 1)
{
    interval_signal = signal(p[interval] to p[interval+1]);
    normalize_interval_x_values between 0 and 1;
    interpolation_grid = from 0 -> 1 in intervallength steps;
    interpolate (interval_x_values, interval_signal)
    given interpolation_grid to obtain interval_signal_interp;
    interpolated_signal[interval] = interval_signal_interp;
}

```

## 2 Activity landscapes fMRI

### **Figure 2.** Pattern based alignment procedure

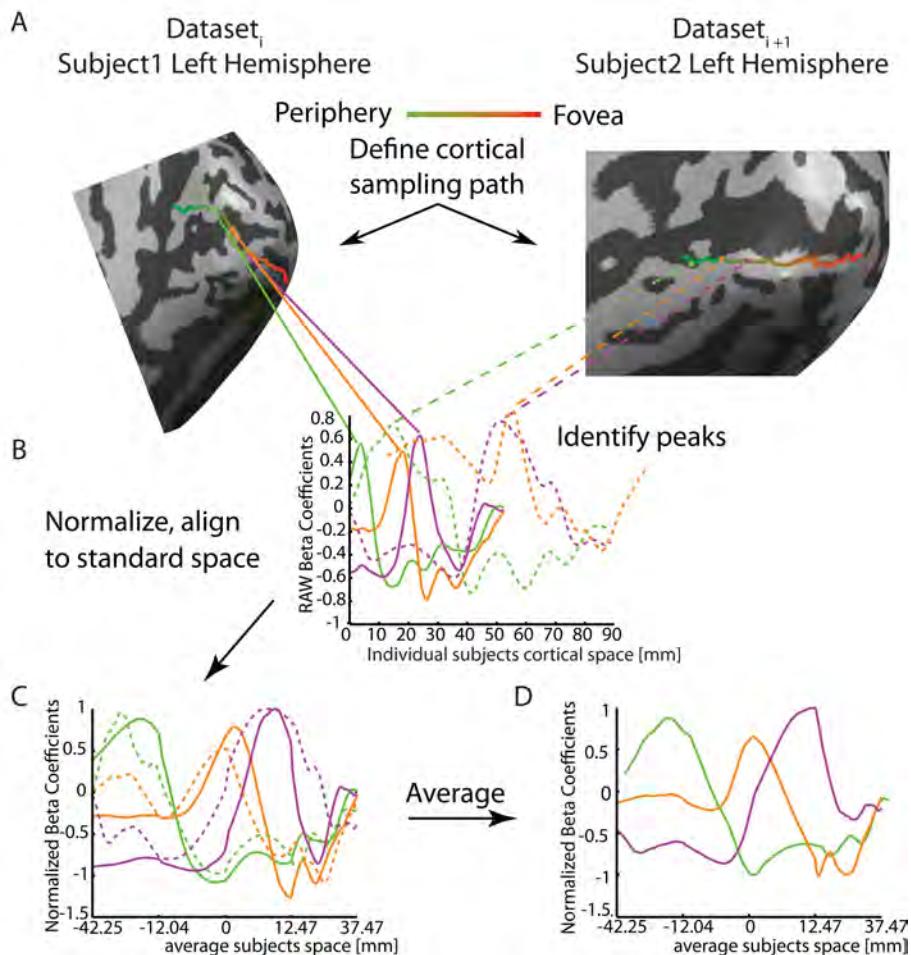
A) Piecewise normalisation and interpolation of spatial data sampled across a mesh path sampled in V1 from fovea to periphery (see Figure 3A for sample paths). In the upper graph pane a signal is shown with three peaks from within-ROI localizers, used as landmarks in the alignment procedure. The lower graph pane shows the gridded signal of interest after normalisation. After normalisation the distance between peaks is constant. B) Pseudocode of the cross PBA algorithm. The signal is split up in four intervals based on three landmark peaks in the signal. Normalization was carried out by dividing segments in between landmark peaks into equal numbers of points ( $N=20$ ), and then standardizing the size of intervals between points. This results in a transformation of the functional data onto a regularly spaced grid (the alignment template).

illustrates the process of aligning and combining sections of activity profiles from multiple cortical paths (in this case from two paths in V1 from two hemispheres in a single individual).

To summarize: to align functional activation along different cortical paths, four steps were required: (i) the definition of the cortical paths; (ii) the determination of the alignment landmarks on the cortical path; (iii), a piecewise transformation of all the intervals between the landmark points, resulting in equalized numbers of standardized intervals and an alignment template; (iv) averaging of functional data from different cortical paths transformed to fit the alignment template.

### **Computing a mesh to mesh mapping**

The one-dimensional PBA described in the previous section can be extended to two dimensions (2D). In this case, the goal may be to align an entire retinotopic map of an area (e.g. V1) across participants and hemispheres, and therefore the entire map can be considered a ROI. The full representation of the cortical sheet after segmentation corresponding to the visual area is represented by a mesh, which is a set of vertices describing the grey-white matter boundary of an anatomical scan connected by triangles (Dale et al., 1999). The mesh used for alignment is isolated from a full brain mesh, flattened, and then submitted to a PBA algorithm as illustrated in Figure 4A-B. The alignment is based on landmark activity from objective localizer stimuli overlaid onto the mesh. We will assume here that the localizer stimuli (e.g., presented at a set of fixed coordinates across participants) provides a 2D pattern of multiple landmarks across the mesh. To compare and align



**Figure 3.** Pattern Based Alignment: illustration of the intermediate steps A) For two left hemispheres (LHs) a cortical path was defined for the V1 region from fovea (red) to periphery (green). B) The signal along the cortical path was plotted (solid lines subject1 LH; dashed lines subject2 LH). Peaks were selected for the different localizers. The cortical location of the peaks is shown in A by dots on the cortical paths. C) After PBA the peaks of the localizers was brought into register. D) Average of the separate localizer signals for subject 1 and subject 2.

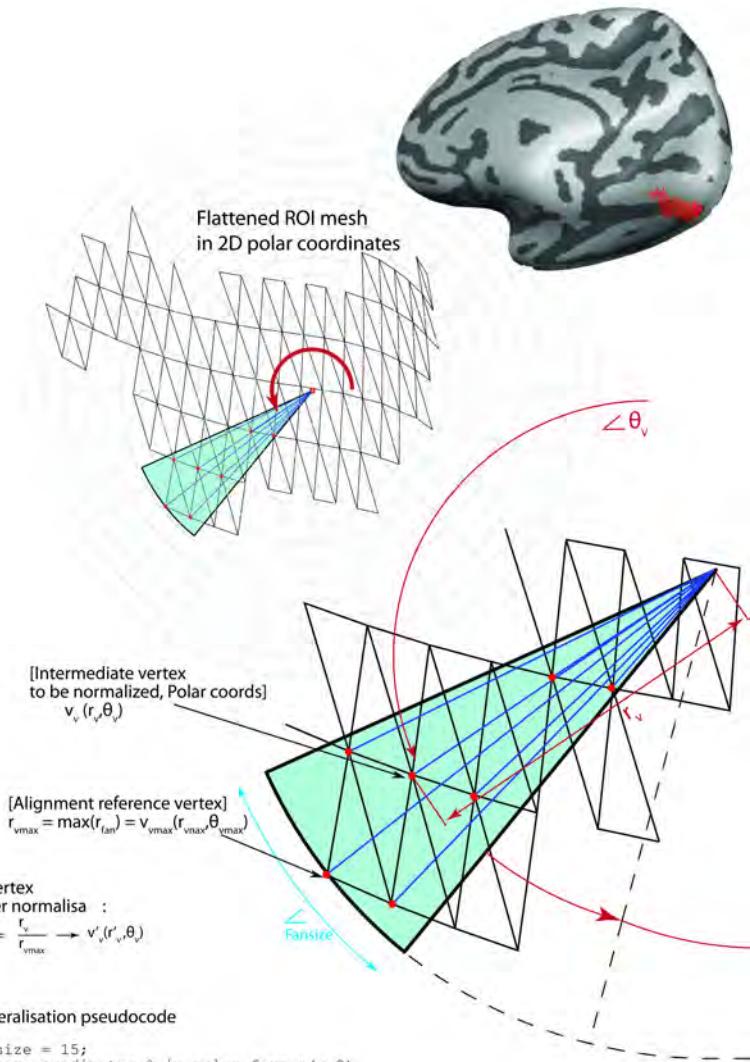
## 2 Activity landscapes fMRI

this activity patterns across different participants, the alignment landmarks have to be normalized into a standardized space. In brief, the PBA algorithm to transform a flattened ROI (a map, or a ROI in the map) into a standardized space comprises four steps: (i) Conversion of the Cartesian coordinates of vertices in the ROI to polar coordinates relative to the centre of the mesh. (ii) Dividing the vertices into fans around the centre point of the ROI. A fan can be visualized as a pie shaped wedge revolving around the centre point of the mesh. At each step around its revolution, vertices falling within the fan are grouped. The size of the fan has to be set based on the number of vertices in the mesh. For example, a sparse mesh needs a larger fan size because a minimal number of vertices needs to fall within the fan at each step around its revolution. (iii) Based on the grouping in the previous step, the radial coordinates of the ROI vertices falling within a fan are normalized through division by the largest radial coordinate within the fan. For example, the radius of a vertex ( $r_v$ ) can be normalized by dividing it with the radius of the vertex with the largest radius ( $r_{v_{max}}$ ), resulting in  $r'_v$ . The angular coordinate of the transformed intermediate vertex is kept  $\Theta_v$  in order to maintain the same position relative to the other intermediate vertices. After steps (ii) and (iii), the new normalized vertices have become the new units of the transformed mesh, characterized by normalized polar coordinates (Figure 4A). (iv) Finally, the still irregularly spaced normalized mesh is then projected and interpolated onto a 2D regular grid.

### Functional alignment of a localizer experiment: comparison between PBA and CBA

We scanned 2 subjects in a visual eccentricity mapping experiment to test the performance of PBA using empirical data. To create landmarks for functional alignment, subjects were instructed to fixate while narrow annuli were presented around fixation. Three annulus sizes were used covering the visual field from fovea to periphery (Figure 5A; for details see experimental procedures). Therefore, along a cortical path from fovea to periphery, we observed three activity peaks corresponding to the three annuli. We opted to demonstrate PBA in V2v data as we found large variability in inter-peak distances in V2v among the 4 different hemispheres used in this experiment. The three peaks served as landmarks for PBA and the resulting aligned cortical path signals were averaged (see Figure 5B). As a comparison we also defined cortical paths through V2v of the down sampled meshes used for CBA. These paths were then CBA transformed and thus projected in

A



B

Generalisation pseudocode

```

fansize = 15;
vertex_coordinates % in polar form v(r,θ)

for(fan=0; fan<360-fansize; fan +=fansize)
{
    make list of vertices within fan
    θ between fan and (fan+fansize);

    find the vertex with the largest radius rvmax

    for(vertex=0; vertex==n_vertices_within_fan ; vertex += 1)
    {
        rv = vertices_within_fan(vertex,1); %get vertex radius
        θv = vertices_within_fan(vertex,2); %get vertex angle

        transformed_vertices_within_fan(vertex,1) = rv/rvmax;
        transformed_vertices_within_fan(vertex,2) = θv;
    }
}

```

## 2 Activity landscapes fMRI

**Figure 4.** A) ROI normalisation procedure. To be able to quantitatively compare activity patterns across different ROIs (based on an objective localizer) they have to be registered into a common space. The algorithm to accomplish this comprises several steps: (i) Conversion of the Cartesian coordinates of the flattened ROI into polar coordinates, (ii) dividing the vertices into fans around the centre point of the ROI, (iii) normalizing the radial coordinate within each fan. (iv) Interpolation of the normalized ROI mesh onto a regular grid.

B) Pseudocode implementation of the ROI normalisation algorithm. A fan rotates in steps around the centre of gravity of the mesh, vertices within the fan are normalized with respect to the vertex with the greatest distance to the mesh centre of gravity.

the average brain space in which the individual hemispheres were aligned. This resulted in the average signal shown in Figure 5C.

Based on the much better preservation of the pattern of functional activity elicited by the landmark activity in PBA compared to CBA, it can be expected that data patterns from independent experiments that are averaged after preceding PBA will be preserved much better than after preceding CBA. We have tested this prediction by applying the two different alignment methods to independent experimental data collected in the same two participants. In this experiment, participants were viewing a centrally presented disk, and a peripherally presented wide annulus. Both modulated temporally in luminance, and were separated by a constant black region in the middle (Figure 5D, for details see experimental procedures). The experimental data obtained in response to the stimulus in Figure 5D were submitted to a GLM with the luminance changes as a predictor. We then applied the alignment template derived from prior PBA of the localizer stimuli to the beta coefficients along the pre-defined anatomical paths yielded by the GLM. This resulted in the pattern of functional data along the cortical path shown in Figure 5E. These aligned data show high Beta coefficients in the representations of the inner disk and outer ring where de luminance change occurred, while Beta coefficients dropped towards zero in the representation of the middle ring where no luminance changes were present. By contrast, alignment and averaging of the experimental data using CBA led to a complete loss of the expected functional data pattern, and averaged Beta values hovered around zero along the entire length of the predefined anatomical path, except for a single, spurious peak in activation (Figure 5F).

This comparison of CBA and PBA shows that PBA is highly efficient in preserving patterns of functional data that are completely lost after CBA. The superior performance of PBA was obtained after comparing the alignment of PBA with that after CBA in only 4 hemispheres. More data would only reinforce the loss of signal after CBA and make the signal after PBA more robust. Thus the example on only four hemispheres used here can serve as a proof of concept for PBA.

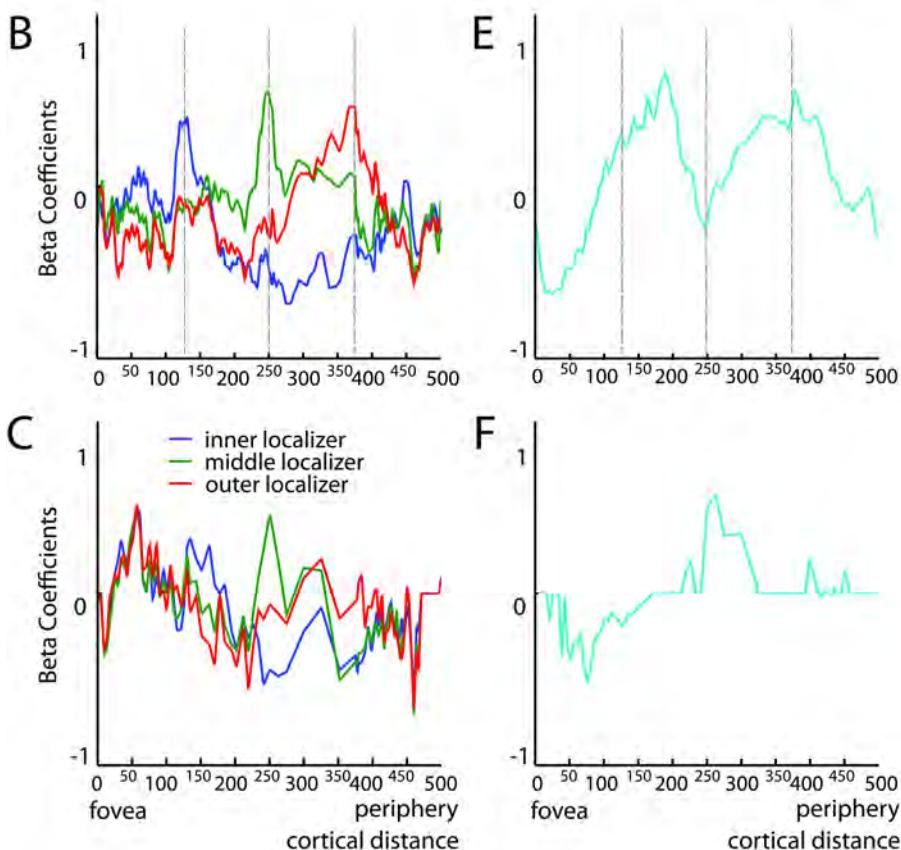
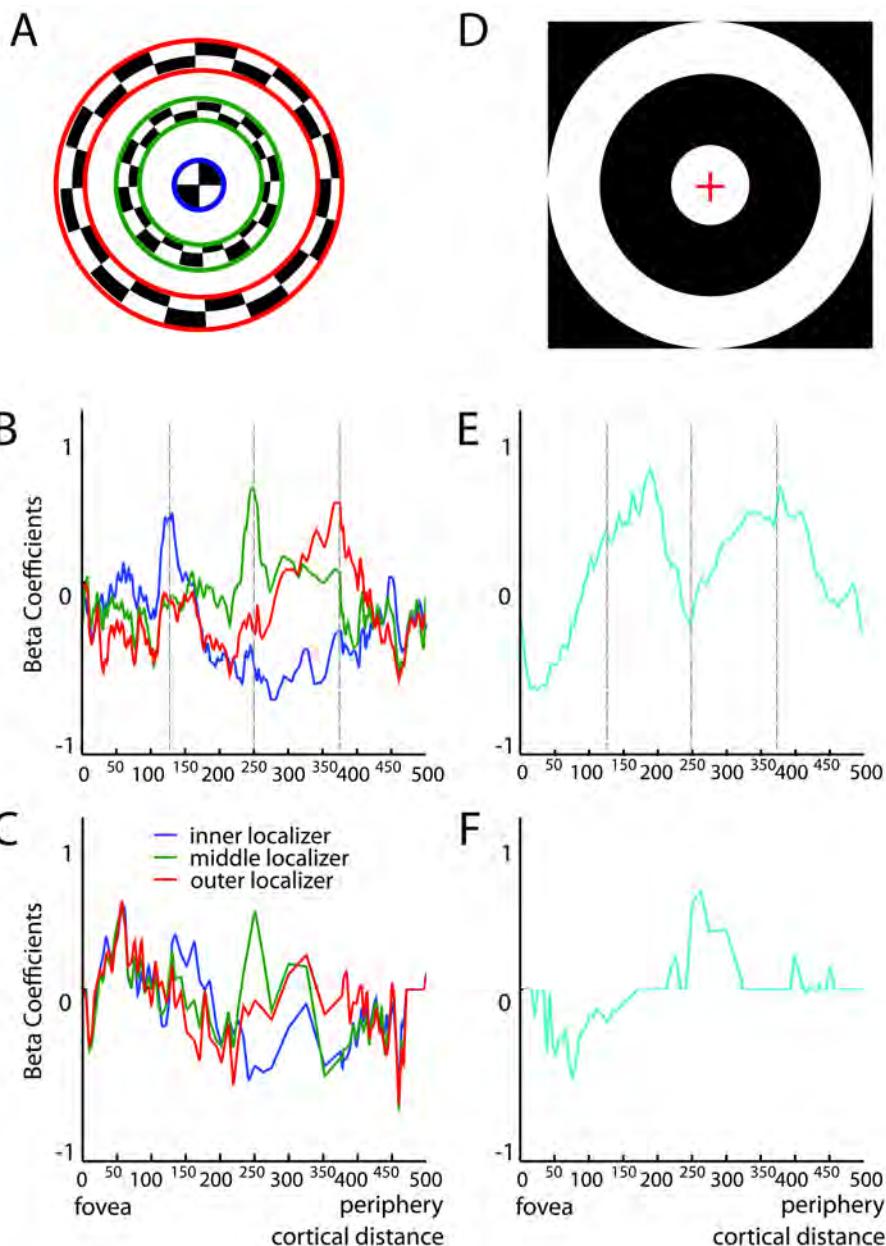
### **Functional alignment of simulated data: comparison between PBA and CBA**

To evaluate the spatial detail that can be resolved by using PBA, we simulated a dataset containing several frequency components (for details see experimental methods, simulations section). In Figure 6, the same simulated dataset was subjected to PBA and CBA, and a quantitative comparison of the quality of functional alignment was carried out.

The synthesized fMRI signal consisted of a sum of sine waves. The overall form of the simulated signal was characterized by 3 peaks, and detail was added to this fundamental signal by 3 harmonics. This signal was enveloped by a Gaussian to create a drop off near the borders of the signal (Figure 6A). The simulated data were generated in a mesh with equally spaced vertices, and were mapped onto the vertices of a subject's native V3 map, using PBA.

To illustrate the effect of CBA and PBA on the aggregation of the simulated datasets, we defined a cortical path through the simulated data (black line in Figure 6A). The data sampled along this cortical path is shown in Figure 6B. A down sampled version of the simulated signal obtained from the CBA mesh is shown in Figure 6C. Cortical paths were defined through both the native (black line in Figure 6D) and the CBA mesh. The data sampled along these paths are shown in Figure 6E after PBA alignment and Figure 6F after reading out the data on the CBA mesh. A comparison between the signal to noise (SNR) of the signal after alignment by PBA ( $\text{SNR}=3.1975 \text{ dB}$ ) and CBA ( $\text{SNR}=-64.4333 \text{ dB}$ ) showed that the signal was well conserved after PBA. After CBA alignment the signal was unidentifiable. To quantify the preservation of fine spatial detail in activity patterns of the PBA and CBA alignment techniques, we analysed the Fourier transform and semblance (Cooper and Cowan, 2008) of the aligned signals. The Fourier transform shows the overall power in the frequency bands of the signal along the cortical path as a whole. The semblance measure shows

## 2 Activity landscapes fMRI



**Figure 5.** Localiser experiment after CBA and PBA

- A) Three annuli centred around a central fixation spot were presented block wise (see experimental procedures) to 2 subjects whom maintained steady fixation.
- B) Cortical paths were defined through V2 in parallel to the horizontal meridian (derived from retinotopy) for the left –and right native resolution mesh of the subject's hemispheres. After identification of the landmarks (vertical gray dash lines), creation of an alignment template, normalisation and averaging, beta coefficient were plotted as a function of cortical space (along the cortical path).
- C) Cortical paths were defined through V2 of the CBA down sampled subject meshes. The cortical paths were transformed into a common brain space by applying CBA. The cortical paths of the 4 hemispheres were then averaged.
- D) The experimental stimulus was made out of two annuli around a disk. The white parts in the Figure showed luminance modulations from white to black (duration of 1s) and back. The two annuli and disk were localized by the checkerboard stimuli in A.
- E) The alignment template obtained by alignment of the localizer peaks (C) was applied to the data of the experimental stimulus and plotted as a function of cortical distance. F) Beta coefficients maps of the experimental stimulus plotted as a function of cortical distance after CBA.

the correspondence in the frequency bands between the reference and the sampled signal at each point along the cortical path.

Figure 6G shows the Fourier power spectrum of the reference signal (Figure 6B, red line), the native PBA signal (Figure 6E, blue line) and the CBA signal (Figure 6F, green line). The power spectrum of the reference signal shows the 4 frequency peaks in the simulated signal. The signal after PBA shows a response to the first 3 peaks, while the signal after CBA only shows a response to the very lowest frequency peak. This indicates that fine grained frequency information was best preserved after PBA alignment. Importantly, this analysis underestimates the efficiency of PBA, because local misalignments distort the Fourier power spectrum.

The interpretation of the semblance results (Figure 6H-I) is similar to a correlation coefficient; it also ranges from minus one to one, with one meaning the best preservation of the original signal. Figure 6H displays the semblance between the reference signal (Figure 6B) and the native sampled signal (Figure 6E), and shows a very high correspondence for low to high frequencies between the PBA aligned signal and the reference signal. The

## *2 Activity landscapes fMRI*

poor semblance for the CBA transformed signal (Figure 6C versus Figure 6F) is shown in Figure 6I. Sampling a limited number of points along the cortical path without PBA resulted in a poor alignment. Only very low spatial frequencies were aligned properly while middle and higher frequency bands lacked semblance between the reference and the sampled CBA signal.

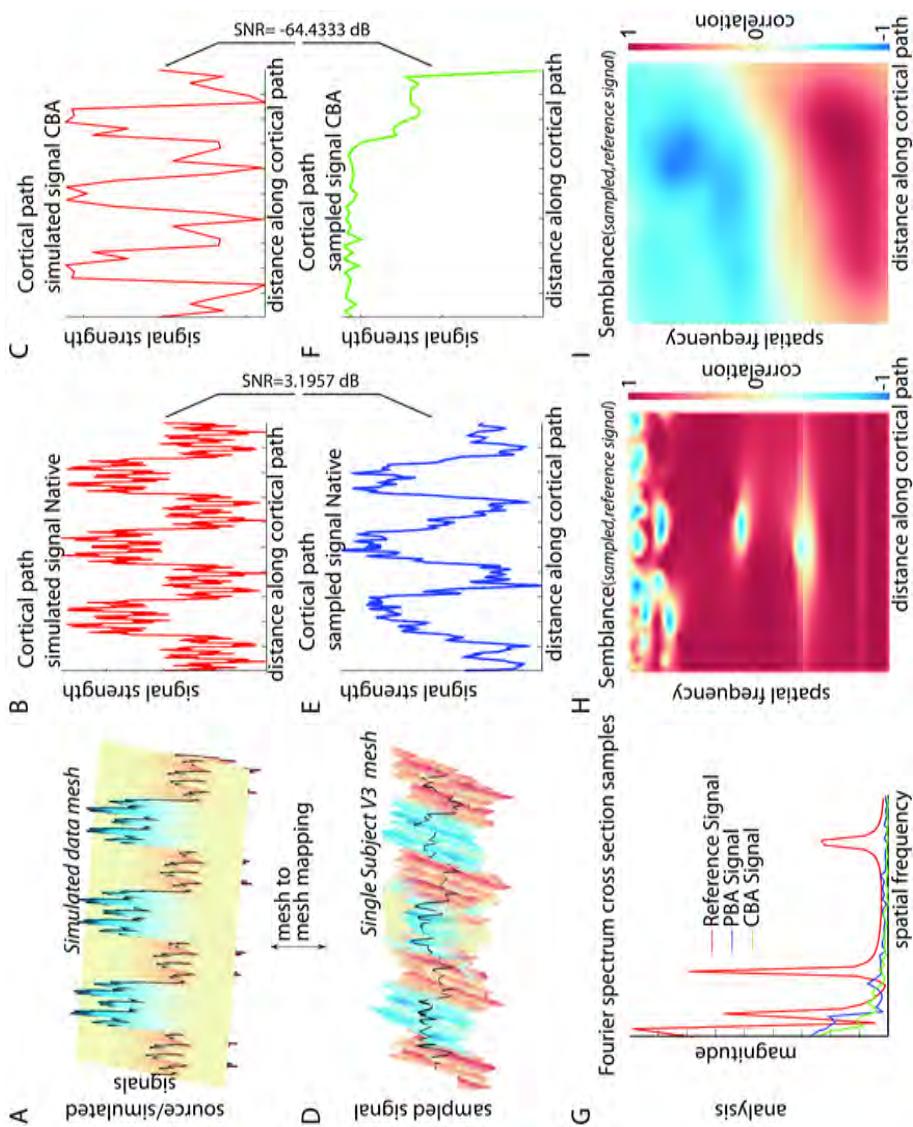
## **Discussion**

This paper demonstrates an alignment technique (PBA) that makes it possible to perform a group analysis of fine-grained, spatial (topographical) patterns in functional data. PBA is based on non-linear morphing of anatomical distances between locations activated within a topographical map by multiple functional localizers (landmarks). By evaluating this technique on both empirical and simulated data after CBA and PBA, we clearly showed that PBA is superior in preserving the fine details of a functional signal after alignment.

PBA can be particularly useful in the evolving field of fMRI research. On the one hand, fMRI paradigms often aim to understand subtle aspects of various types of sensory processing and cognition. For example, a number of studies have aimed to find fMRI correlates of visual illusions to study the processes underlying boundary and surface perception (Komatsu, 2006; Goebel and De Weerd, 2009). These paradigms often require the analysis of small signals that are lost without optimal alignment. On the other hand, high field fMRI permits (or is close to permitting) imaging at the level of columns and layers (Cheng et al., 2001; Kim and Ogawa, 2002; Silva and Koretsky, 2002; Duyn et al., 2007). To preserve patterns of activity distributed over these cortical compartments after alignment across participants, new alignment techniques such as PBA are extremely useful.

## **Functional alignment in topographic maps using PBA**

In a recent study, we used PBA to reveal a small signal related to brightness induction (Van de Ven et al., submitted manuscript). The area within which the induction was generated consisted of an annulus with inner and outer radii of 3° and 9° respectively. The luminance of this area was constant (grey), while in the disk enclosed within the grey annulus and in a large annulus enclosing the grey annulus (the inducers), the luminance was modulated dynamically. The luminance modulation in the inducers generated an antiphase brightness percept in the grey region of constant luminance,



## 2 Activity landscapes fMRI

**Figure 6.** Sensitivity simulations of mesh mapping method.

- A) A composite sinusoidal signal consisting of 4 frequencies enveloped by a Gaussian with a periodicity that allowed for 3 peaks was synthesized in order to be projected on a mesh region corresponding to a visual area. The solid black line shows the sampling path on which the PBA was based.
- B) A reference signal for PBA was obtained by reading out the signal along the cortical path in (a).
- C) A reference signal for CBA was obtained by down sampling the reference signal in (b) (with respect to the vertex distance of a CBA mesh).
- D) The synthesized signal in (a) was projected onto V2v hemisphere. The black solid line shows the path parallel to the horizontal meridian for PBA.
- E) The signal along the sampling paths (average of 4 hemispheres of 2 subjects) in (d) shows the three peaks as well as some of the higher frequency information present in the synthesized signal (a).
- F) Signal of the simulated data when projected onto a CBA mesh (average of 4 hemispheres of 2 subjects).
- G) Fourier spectra of the reference (subfigure B; red solid line), the sampled signal in the native resolution (subfigure E; blue solid line) –and the sampled signal in the CBA resolution (subfigure F; green solid line).
- H) Semblance plot of the sampled signal (e) and the reference signal (b), both signals are decomposed into frequency components over cortical distance, semblance shows the degree of correspondence between the 2 signals for all frequencies at specific points along the cortical path. Colors ranging from blue (-1) to red (+1) code the correlation coefficient, depict the resemblance, of the 2 signals in the frequency space domain.
- I) Semblance of the signal after CBA (f) and the reference signal for CBA (c). Color conventions are the same as in (h).

which did not occur when the region of constant luminance was black. In agreement with prior observations in single cell recording studies (Rossi et al., 1996; Rossi and Paradiso, 2003), the signal corresponding to this illusion was very weak, and was undetectable without PBA. We carried out PBA along a cortical path through the representations of inducers and an interspersed area of constant luminance in V1, V2 and V3, using peak activity of localizers centred in each stimulus part as landmarks. The same transformations were then applied to the functional datasets from the induction experiment. This helped us to reveal antiphase activity in the region of constant luminance in V2, when that region was grey, but not when it was black. Thus, PBA was instrumental in this study in demonstrating a corre-

late of surface perception in early visual cortex, thereby contributing to an important debate in visual neuroscience (Cornelissen et al., 2006; Komatsu, 2006). Using PBA, we also recently demonstrated for the first time an fMRI correlate of texture filling-in in early visual cortex in a Troxler fading paradigm (Jans et al., submitted manuscript), replicating earlier findings from single cell recordings in monkeys (De Weerd et al., 1995).

Another important debate in visual neuroscience is the extent to which deafferentation (due to a retinal scotoma) leads to topographical remapping in early visual cortex. A long tradition of neurophysiological (for reviews see: Buonomano and Merzenich, 1998; Kaas et al., 2003) and molecular work (Arckens et al., 2000; Arckens et al., 2003; Kaas et al., 2003) has accumulated evidence for retinotopic re-mapping, during which neurons in the deafferented cortical region gain receptive fields outside the scotoma corresponding to the deafferented region. However, more recent work combining neurophysiology and high-field neuroimaging in monkeys (Smirnakis et al., 2005) has put in doubt the extent of this reorganization. While it is likely that the reorganization is less important (or applies to smaller numbers of cells) than hitherto assumed, it seems unlikely that there is no reorganization at all. We suggest that PBA of fMRI data in large groups may increase sensitivity to detect these plastic changes.

In addition to studying sensory processes, PBA can be useful to enhance sensitivity for the modulation of sensory processes by high-level cognitive factors, such as short-term memory and attention. In the field of attention, an important unresolved question is whether spatial attention can be divided into multiple foci (Jans et al., 2010). If attention could be divided in multiple foci, this will likely entail a cost (Cave et al., 2010) so that the attentional enhancement of visual processing in retinotopic regions corresponding to the hypothesized foci is likely to be small. In these experiments, and others using subtle cognitive manipulations of visual processing in retinotopic maps, we suggest that PBA can be an important tool to enhance sensitivity for fMRI correlates. Importantly, while we have stressed the potential of PBA in visual neuroscience, this method of alignment can be applied equally successfully in all other sensory modalities showing topographical representations of sensory space. For example, PBA can be useful in somatosensory, motor and auditory investigations of topographic changes in conditions of use (training) or under-use (deafferentations) (for reviews see: Buonomano and Merzenich, 1998; Pessoa and De Weerd, 2003), and investigations that monitor and influence adaptive and maladaptive effects of plasticity in topographic maps (Ziemann, 2004).

## **Applications of PBA in high field imaging**

Since the advent of high-field and ultra-high field fMRI the spatial resolution of the acquired functional and structural images has increased to the point that functional imaging at the level of resolution of columns and layers is within reach (Chen and Ugurbil, 1999; Kim and Ogawa, 2002; Logothetis, 2008). Several fMRI studies have revealed in animals and humans the underlying columnar organisation of cortex (Ohki et al., 2006; Moon et al., 2007; Ohki and Reid, 2007; Yacoub et al., 2007; Yacoub et al., 2008), and others have demonstrated in humans the existence of ocular dominance domains and orientation columns in primary visual cortex (Ohki et al., 2006; Yacoub et al., 2007; Yacoub et al., 2008). These functional details in 2D topographic maps can be exploited to bring topographic maps of entire cortical areas in register across hemispheres and participants. For example, when identified, the centre of a pinwheel representing an orientation column could serve as an alignment landmark. The centres of all the imaged pinwheels together form an irregularly spaced grid which can be aligned between areas or subjects. Furthermore, alignment by matching pinwheel centers would open the possibility to separate processes in blobs and interblobs, and investigate different functions of these anatomical compartments in sensory function (Livingstone and Hubel, 1983; Roe and Ts'o, 1999; Sincich and Horton, 2005). Detailed landmarks in topographic maps cannot only help to bring functional maps of different hemispheres and participants in alignment, they also permit direct mapping of model data onto meshes containing the empirical data. This, in turn, permits direct tests of model predictions against empirical data in the brain (Goebel and De Weerd, 2009; Peters et al., 2010). Moreover, thanks to the possibility to image the laminae of the cortical sheet (Harel et al., 2006), separate mesh models of, for example, granular, infra-granular, and supra-granular layers can be constructed, which combined with functional connectivity analysis (Roebroeck et al., 2005) might provide a window on detailed feedforward – feedback interactions. Again, this would critically depend on accurate alignment of maps within and across participants, where PBA can be useful.

## **Conclusion**

We have shown a pattern based alignment technique that makes it possible to perform a group analysis of fine-grained, spatial (topographical) patterns in functional data, and how this method can be useful in existing paradigms as well as in the evolving field of high field imaging.

## References

- Aguirre, G.K., Zarahn, E., and D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron* 21, 373-383.
- Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T., and Zilles, K. (2000). Brodmann's areas 17 and 18 brought into stereotaxic space-where and how variable? *Neuroimage* 11, 66-84.
- Arckens, L., Van der Gucht, E., Van den Bergh, G., Massie, A., Leysen, I., Vandenbussche, E., Eysel, U.T., Huybrechts, R., and Vandesande, F. (2003). Differential display implicates cyclophilin A in adult cortical plasticity. *Eur J Neurosci* 18, 61-75.
- Arckens, L., Schweigart, G., Qu, Y., Wouters, G., Pow, D.V., Vandesande, F., Eysel, U.T., and Orban, G.A. (2000). Cooperative changes in GABA, glutamate and activity levels: the missing link in cortical plasticity. *Eur J Neurosci* 12, 4222-4232.
- Brett, M., Johnsrude, I.S., and Owen, A.M. (2002). The problem of functional localization in the human brain. *Nature Reviews Neuroscience* 3, 243-249. Brindley, G.S. (1972). The variability of the human striate cortex. *Journal of Physiology* 225, 1P-3P.
- Buonomano, D.V., and Merzenich, M.M. (1998). Cortical plasticity: from synapses to maps. *Annu Rev Neurosci* 21, 149-186.
- Cave, K.R., Bush, W.S., and Taylor, T.G. (2010). Split attention as part of a flexible attentional system for complex scenes: comment on Jans, Peters, and De Weerd (2010). *Psychol Rev* 117, 685-696.
- Chau, W., and McIntosh, A.R. (2005). The Talairach coordinate of a point in the MNI space: how to interpret it. *Neuroimage* 25, 408-416.
- Chen, W., and Ugurbil, K. (1999). High spatial resolution functional magnetic resonance imaging at very-high-magnetic field. *Top Magn Reson Imaging* 10, 63-78.
- Cheng, K., Waggoner, R.A., and Tanaka, K. (2001). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron* 32, 359-374.
- Clarke, S., and Miklossy, J. (1990). Occipital cortex in man: organization of callosal connections, related myelo- and cytoarchitecture, and putative boundaries of functional visual areas. *Journal Of Comparative Neurology* 298, 188-214.
- Cooper, G.R.J., and Cowan, D.R. (2008). Comparing Time Series using Wavelet Based Semblance Analysis. *Computers & Geosciences* 34, 95-102.
- Cornelissen, F.W., Wade, A.R., Vladusich, T., Dougherty, R.F., and Wandell, B.A. (2006). No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex. *J Neurosci* 26, 3634-3641.
- Dale, A.M., Fischl, B., and Sereno, M.I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9, 179-194. Daniel, P.M., and Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *J Physiol* 159, 203-221.
- Datta, R., and DeYoe, E.A. (2009). I know where you are secretly attending! The

## *2 Activity landscapes fMRI*

- topography of human visual attention revealed with fMRI. *Vision Research* 49, 1037-1044.
- Dougherty, R.F., Koch, V.M., Brewer, A.A., Fischer, B., Modersitzki, J., and Wandell, B.A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal Of Vision* 3, 586-598.
- Dumoulin, S.O., and Wandell, B.A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage* 39, 647-660.
- Duyn, J.H., van Gelderen, P., Li, T.Q., de Zwart, J.A., Koretsky, A.P., and Fukunaga, M. (2007). High-field MRI of brain cortical substructure based on signal phase. *Proc Natl Acad Sci U S A* 104, 11796-11801.
- Engel, S.A., Glover, G.H., and Wandell, B.A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex* 7, 181-192.
- Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B.T., Mohlberg, H., Amunts, K., and Zilles, K. (2008). Cortical folding patterns and predicting cytoarchitecture. *Cereb Cortex* 18, 1973-1980.
- Fischl, B., Sereno, M.I., and Dale, A.M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9, 195-207.
- Friston, K.J., Rotshtein, P., Geng, J.J., Sterzer, P., and Henson, R.N. (2006). A critique of functional localisers. *Neuroimage* 30, 1077-1087. Gattass, R., Gross, C.G., and Sandell, J.H. (1981). Visual topography of V2 in the macaque. *J Comp Neurol* 201, 519-539.
- Goebel, R., Esposito, F., and Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Human Brain Mapping* 27, 392-401.
- Goebel, R., and De Weerd, P. (2009). Perceptual Filling-in: From Experimental Data to Neural Network Modeling. In *The Cognitive Neurosciences*, M.S. Gazzaniga, ed. (Cambridge, MS.: The MIT Press), pp. 435-454.
- Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Res* 41, 1409-1422.
- Harel, N., Lin, J., Moeller, S., Ugurbil, K., and Yacoub, E. (2006). Combined imaging-histological study of cortical laminar specificity of fMRI signals. *Neuroimage* 29, 879-887.
- Holmes, G. (1919). The cortical localization of vision. In *The Montgomery Lectures in Ophthalmology* (Trinity College Dublin). Horton, J.C., and Hoyt, W.F. (1991). The representation of the visual field in human striate cortex. A revision of the classic Holmes map. *Archives Of Ophthalmology* 109, 816-824.
- Ishai, A., Ungerleider, L.G., Martin, A., Schouten, J.L., and Haxby, J.V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proc Natl*

## References

- Acad Sci U S A 96, 9379-9384.
- Jans, B., Peters, J.C., and De Weerd, P. (2010). Visual spatial attention to multiple locations at once: the jury is still out. *Psychol Rev* 117, 637-684.
- Kaas, J.H., Collins, C.E., and Chino, Y.M. (2003). The Reactivation and Reorganization of Retinotopic Maps in Visual Crotex of Adult Mammals After Retinal and Cortical Lesions. In *Filling-in: From Perceptual Completion to Cortical Reorganization*, L. Pessoa, and P. De Weerd, eds. (New York: Oxford University Press).
- Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17, 4302-4311.
- Kim, S.G., and Ogawa, S. (2002). Insights into new techniques for high resolution functional MRI. *Curr Opin Neurobiol* 12, 607-615.
- Kirson, D., Huk, A.C., and Cormack, L.K. (2008). Quantifying spatial uncertainty of visual area boundaries in neuroimaging data. *Journal Of Vision* 8, 10 11-15.
- Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nat Rev Neurosci* 7, 220-231.
- Lancaster, J.L., Tordesillas-Gutierrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J.C., and Fox, P.T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum Brain Mapp* 28, 1194-1205.
- Livingstone, M.S., and Hubel, D.H. (1983). Specificity of cortico-cortical connections in monkey visual system. *Nature* 304, 531-534.
- Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. *Nature* 453, 869-878.
- Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A* 92, 8135-8139.
- Mitsis, G.D., Iannetti, G.D., Smart, T.S., Tracey, I., and Wise, R.G. (2008). Regions of interest analysis in pharmacological fMRI: how do the definition criteria influence the inferred result? *Neuroimage* 40, 121-132.
- Moon, C.H., Fukuda, M., Park, S.H., and Kim, S.G. (2007). Neural interpretation of blood oxygenation level-dependent fMRI maps at submillimeter columnar resolution. *J Neurosci* 27, 6892-6902.
- Ohki, K., Chung, S., Kara, P., Hubener, M., Bonhoeffer, T., and Reid, R.C. (2006). Highly ordered arrangement of single neurons in orientation pinwheels. *Nature* 442, 925-928.
- Ohki, K., and Reid, R.C. (2007). Specificity and randomness in the visual cortex. *Curr Opin Neurobiol* 17, 401-407.
- Ono, M., Kubik, S., and Abernathy, C.D. (1990). *Atlas of the Cerebral Sulci* (New York:

## *2 Activity landscapes fMRI*

Thieme Medical).

- Pessoa, L., and De Weerd, P. (2003). *Filling-in: From Perceptual Completion to Cortical Reorganization* (New York: Oxford University Press).
- Peters, J.C., Jans, B., van de Ven, V., De Weerd, P., and Goebel, R. (2010). Dynamic brightness induction in V1: Analyzing simulated and empirically acquired fMRI data in a "common brain space" framework. *Neuroimage* 52, 973-984.
- Qiu, A., Rosenau, B.J., Greenberg, A.S., Hurdal, M.K., Barta, P., Yantis, S., and Miller, M.I. (2006). Estimating linear cortical magnification in human primary visual cortex via dynamic programming. *Neuroimage* 31, 125-138.
- Roe, A.W., and Ts'o, D.Y. (1999). Specificity of color connectivity between primate V1 and V2. *J Neurophysiol* 82, 2719-2730.
- Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage* 25, 230-242.
- Rossi, A.F., and Paradiso, M.A. (2003). Surface Completion: Psychophysical and Neurophysiological Studies of Brightness. In *Filling-in: From Perceptual Completion to Cortical Reorganization*, L. Pessoa, and P. De Weerd, eds. (New York: Oxford University Press).
- Rossi, A.F., Rittenhouse, C.D., and Paradiso, M.A. (1996). The representation of brightness in primary visual cortex. *Science* 273, 1104-1107.
- Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889-893.
- Sergent, J., Ohta, S., and MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain* 115 Pt 1, 15-36.
- Silva, A.C., and Koretsky, A.P. (2002). Laminar specificity of functional MRI onset times during somatosensory stimulation in rat. *Proc Natl Acad Sci U S A* 99, 15182-15187.
- Sincich, L.C., and Horton, J.C. (2005). Input to V2 thin stripes arises from V1 cytochrome oxidase patches. *J Neurosci* 25, 10087-10093.
- Smirnakis, S.M., Brewer, A.A., Schmid, M.C., Tolias, A.S., Schuz, A., Augath, M., Inhoffen, W., Wandell, B.A., and Logothetis, N.K. (2005). Lack of long-term cortical reorganization after macaque retinal lesions. *Nature* 435, 300-307.
- Stensaas, S.S., Eddington, D.K., and Dobelle, W.H. (1974). The topography and variability of the primary visual cortex in man. *Journal Of Neurosurgery* 40, 747-755.
- Talairach, J., and Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-Dimensional proportional system - An approach to cerebral imaging* (New York: Thieme).
- Watson, J.D., Myers, R., Frackowiak, R.S., Hajnal, J.V., Woods, R.P., Mazziotta, J.C., Shipp, S., and Zeki, S. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb*

## *References*

- Cortex 3, 79-94.
- Wilms, M., Eickhoff, S.B., Homke, L., Rottschy, C., Kujovic, M., Amunts, K., and Fink, G.R. (2010). Comparison of functional and cytoarchitectonic maps of human visual areas V1, V2, V3d, V3v, and V4(v). *Neuroimage* 49, 1171-1179.
- Yacoub, E., Harel, N., and Ugurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proceedings National Academy Science U S A* 105, 10607-10612.
- Yacoub, E., Shmuel, A., Logothetis, N., and Ugurbil, K. (2007). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *Neuroimage* 37, 1161-1177.
- Zeki, S., Watson, J.D., Lueck, C.J., Friston, K.J., Kennard, C., and Frackowiak, R.S. (1991). A direct demonstration of functional specialization in human visual cortex. *J Neurosci* 11, 641-649.
- Ziemann, U. (2004). TMS induced plasticity in human cortex. *Rev Neurosci* 15, 253-266.



### **3 A correlate of brightness induction in human V2**

based on: Van de Ven, V., Jans, B., Goebel, R., De Weerd, P., A correlate of brightness induction in human V2. (accepted)

### *3 Brightness Induction*

## **Abstract**

Visual scene perception owes greatly to surface features such as color and brightness. Yet, early visual cortical areas predominantly encode surface boundaries, rather than surface interiors. Whether human early visual cortex may nevertheless carry a small signal relevant for surface perception is a topic of debate. We induced brightness changes in a physically constant surface by temporally modulating the luminance of surrounding surfaces in seven human participants. We found that functional magnetic resonance imaging (fMRI) activity in the V2 representation of the constant surface was in anti-phase to luminance changes of surrounding surfaces (i.e., activity was in-phase with perceived brightness changes). Moreover, the amplitude of the anti-phase fMRI activity in V2 predicted the strength of illusory brightness perception. We interpret our findings as evidence for a surface-related signal in early visual cortex, and discuss the neural mechanisms that may underlie that signal in concurrence with its possible interaction with properties of the fMRI signal.

## **Introduction**

Traditional views and empirical work of early vision have emphasized the role of local contrast in the reconstruction of the visual image (Hubel & Wiesel, 1962). However, mechanisms extracting local contrast appear ill-suited to encode homogenous surfaces, and despite a significant amount of work, the question by which means surface properties are encoded remains unresolved (Komatsu, 2006; Pessoa, Thompson, & Noe, 1998). Neurophysiological studies have demonstrated that a minority of neurons responds to homogenous surface luminance (Haynes, Lotto, & Rees, 2004; Kinoshita & Komatsu, 2001; Roe, Lu, & Hung, 2005; Rossi, Rittenhouse, & Paradiso, 1996), and fMRI has confirmed the existence of a signal related to surface luminance in early visual cortex (Haynes et al., 2004). Furthermore, neurophysiological studies indicate that there are neurons in anatomical sub-compartments of V2 that carry a signal related to the perception of surface brightness in the Craik-O'Brian Cornsweet Illusion (Roe et al., 2005), and other studies showed a correlate of brightness perception in a small number of V1 neurons using an induced brightness paradigm (Rossi et al., 1996; Rossi and Paradiso, 1999). The goal of the present study was to investigate whether the signal related to surface brightness perception that has

## *Methods*

been reported by neurophysiological and optical imaging studies in cats and monkeys could also be demonstrated with fMRI in humans.

To study the perception of surface brightness, we used a dynamic brightness induction paradigm similar to that used by Paradiso and colleagues, who recorded from neurons in cat area 17 (primary visual cortex) while a grey surface of constant luminance was placed over their receptive fields (RFs) (Rossi et al., 1996). A minority of neurons (approx. 10%) responded in anti-phase to luminance modulations in abutting inducer surfaces far outside the RFs, which suggests that these neurons helped encoding the counter-phase brightness changes that humans perceive in the same stimuli. The illusion and its neural correlate were present with a fixed grey surface (induction condition), and absent with a black surface (control condition). We will refer to the area of fixed luminance as the probing region, as it was designed to probe mechanisms of surface perception.

The anti-phase relationship between perceived brightness and inducer luminance is the hallmark of brightness induction (Rossi et al., 1996). To demonstrate this property in human visual cortex with functional magnetic resonance imaging (fMRI), we used a slow event-related design in which changes between high and low inducer luminance levels were separated by several seconds, thereby taking into account the slow fMRI hemodynamic signal (Boynton, Engel, Glover, & Heeger, 1996). Furthermore, we optimized the spatial dimensions of the stimulus to take into account the spatial resolution limits of the fMRI signal (Cornelissen, Wade, Vladusich, Dougherty, & Wandell, 2006; Sereno et al., 1995) and limits in the spatial extent of brightness induction (Pereverzeva & Murray, 2008). A possible brightness-related anti-phase signal in the probing region could be partly annulled or masked by opposite-phase signals from abutting inducers, which would be strongest near the retinotopic projections of the inducer and probing region borders. Therefore, we hypothesized that a brightness-related anti-phase fMRI signal would be strongest near the middle of the retinotopic projection of the probing region, and would counteract in-phase responses to inducers. We used localizer stimuli to delineate retinotopic projections of inducers and probing region.

## **Methods**

Participants Seven healthy adults (25 - 30y) with (corrected-to) normal visual acuity participated after written informed consent. All participants completed the main fMRI brightness induction and control experiments with

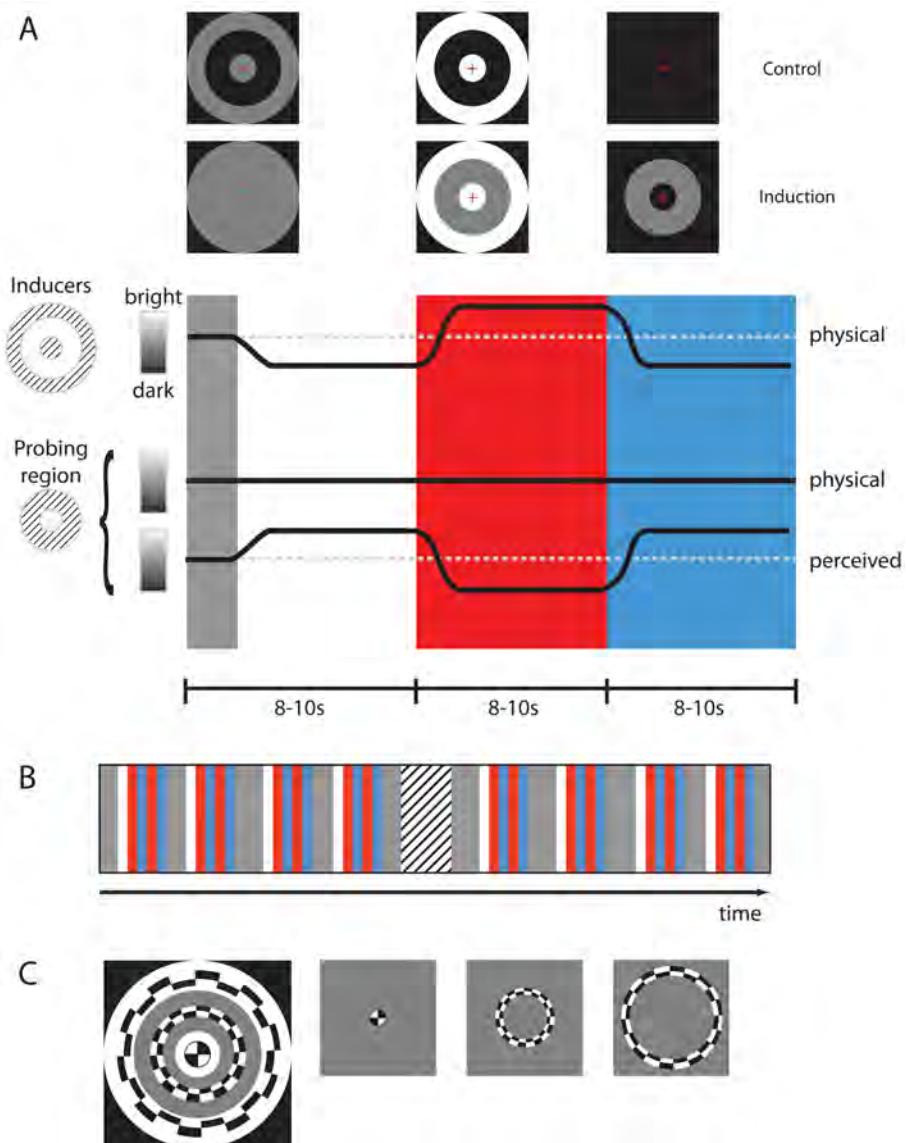
### *3 Brightness Induction*

stimulus patterns that comprised two inducers, as well as associated retinotopy and eccentricity localizers, and brightness psychophysical rating task. Three of the participants completed additional fMRI experiments in which luminance changes were presented in only one inducer, as well as associated psychophysical ratings of brightness induction

### **Stimulus Design**

Radius of the inner inducer disk was 3°, width of the probing region was 6°, and width of the outer inducer annulus was 5°. Stimuli were presented concentrically around a fixation cross (1° width and height) placed on the inner inducer. The probing region was grey ( $7.9 \text{ cd/m}^2$ ; induction condition) or black ( $1.1 \text{ cd/m}^2$ ; control condition). The inducers changed between minimum and maximum luminance in a sinusoidal fashion in 1s (range =  $1.1 - 164 \text{ cd/m}^2$ ) (cf, Rossi and Paradiso, 1999). During inducer luminance modulations, contrast at the borders with the probing region varied between -76% and 91% in the induction condition, and between 0% and 99% in the control condition (Michelson indexes). Figure 1A shows the induction and control stimuli at different luminance levels of the inducers.

The size of the probing region satisfied two opposite demands: maximization of probing region size to reduce effects of fMRI spatial signal spread (Sereno et al., 1995) (De Weerd, P., Karni, A., Kastner, S., Ungerleider, L. G., and Jezzard, P. Presented at the Third International Conference on Functional Mapping of the Human Brain, 1997), and minimization of probing region size to maximize the illusion across its extent (Pereverzeva & Murray, 2008). First, the spatial spread of fMRI signal dictates that the probing region should be maximized because the signals associated with probing region and inducers were expected to be in anti-phase during brightness induction. Anti-phase signals attenuate each other when they spread into each other's retinotopic territory. Point spread of the fMRI signal has been estimated to be on the order of just a few mms (Sereno et al., 1995; Engel et al., 1997], but recent estimates suggested spread of approximately 7mm at half-width-at-half-maximum (HWHM) (Cornelissen et al., 2007). According to an estimate of V1 cortical magnification by (Sereno et al., 1995), the distance from the inner to the outer border of the probing region projected on V1 is about 15mm. In light of this estimate of fMRI signal spread and chosen stimulus dimensions, a sufficient proportion of a hypothetical counter-phase signal related to brightness induction was expected to survive in the middle of the probing region, potentially even directly revealing anti-phase activity (depending on the strength of that signal). Going



### *3 Brightness Induction*

**Figure 1.** Stimuli and experimental design.

(A) Brightness induction and control stimuli (top), and timing of events (bottom). A fixation period (grey area in timing diagram) preceded each block of luminance changes. A block started with a grey-to-dark change in inducers (white area in timing diagram), followed by dark-to-bright (red area) and bright-to-dark changes (blue area) (1s duration) after which the new luminance value was kept constant for a variable inter-trial-interval (ITI 7-9s). The block ended with a dark-to-grey change (not shown). Luminance changes occurred only in the inducers (concentric rings, left inset), while probing region luminance (single annulus inset) was kept constant throughout the measurement (straight line timing diagram). The brightness perception in the grey probing region (bottom timing diagram) follows a course inverse to the physical luminance changes in inducers. (B) Schematic of a single run (color coding as in Figure. 1A). Between the fourth and fifth block a short ‘relaxation period’ was used (black-and-white textured region), during which participants could briefly refrain from fixation (see Supplementary Material – Experimental Design). (C) Localizer rings map onto the center of the probing and inducer regions of the stimuli. Smaller insets illustrate localizer stimulus separately for the central inducer (left), probing region (middle) and peripheral inducer areas (right) (stimuli are scaled for visualization purposes).

from the middle of the probing region to its border, in-phase activity from inducers was expected to increasingly dominate the hypothetical anti-phase signal in the probing region. A hypothetical anti-phase signal associated with induction would also spread into the territory of the inducers, and thereby reduce in-phase signal of inducers. Thus, suppression of in-phase activity from inducers is an important and integral aspect of expected results during brightness induction. We anticipated that smaller cortical magnification and other factors in extrastriate areas could render the separation of signals from probing and inducer regions more difficult. Second, although the spatial spread of fMRI signal argues in favour of making the probing region as large as possible, this manipulation also decreases the strength of the illusion, which puts a limit on the extent to which the probing region can be enlarged. Using a probing region with a cortical projection in V1 that was larger than ours (18mm), Pereverzeva and colleagues (Pereverzeva & Murray, 2008) found a decline of brightness induction towards the middle of the probing region. The smaller probing region in our study (both in the visual field and on cortex) was chosen to limit the decline of the illusion towards the middle.

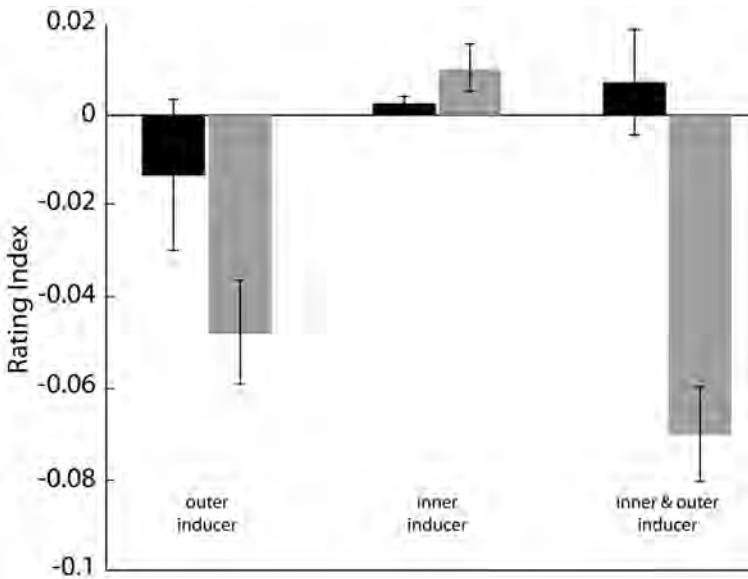
Localizer stimuli were centred in the three parts of the stimuli. The inner radii of central, middle and outer localizer rings were 0°, 5.4° and 11°, respectively, and their outer radii 0.5°, 6.6° and 13°, respectively. Each localizer consisted of 20 sequentially positioned black-and-white segments, which reversed contrast at a rate of 16 Hz. Localizer stimuli were presented on a fixed grey background ( $7.9 \text{ cd/m}^2$ ).

Stimuli were projected onto a diagonally positioned mirror attached to the head coil in the scanner bore using an LCD-projector controlled by a PC (screen refresh rate, 60 Hz) running Presentation software (version 9.90, Neurobehavioral Systems Inc.). The onsets of luminance modulations were pulse-triggered by the T2\*-weighted image acquisition.

### **Psychophysical Ratings of Induction**

All participants rated the strength of induced brightness change in a psychophysical experiment in the scanner. Participants mimicked perceptual changes in the probing region of an experimental stimulus by setting a physical luminance value in the probing region of a testing stimulus. Participants were presented with 6 conditions: Three types of stimulus display (two inducers, only inner inducer, only outer inducer) combined with the two luminance levels of the probing region (grey and black). For each condition, participants completed 24 trials. Each trial comprised 1) a 1s starting phase, in which one or both inducers were shown at their maximum or minimum luminance, 2) a luminance modulation phase, in which one or both inducers gradually changed luminance in 1s from maximum to minimum (down-sweep) or minimum to maximum luminance (up-sweep), 3) a 1s end phase, in which one or both inducers remained at final luminance, and 4) a masking phase, in which the surfaces of both inducers (but not the probing region) were masked by a black-and-white checkerboard mask. At the start of each trial, the probing region was set to grey or black, which remained fixed throughout the four trial phases. During the masking phase, participants could manually set on a continuous scale (using two buttons) the luminance level of the probing region to the luminance level that matched perceived brightness at the end of the modulation phase. The physical luminance of the probing region and the luminance set to match induction ( $\text{cd/cm}^2$ ) were log-transformed (natural log). A rating index (RI) was obtained by subtracting the log-transformed starting luminance values of the probing region from the log-transformed ratings provided by the subjects. The resulting RI was negative when the luminance deviation in probing region was set in anti-phase to inducer luminance modulations, and a positive

### 3 Brightness Induction



**Figure 2.** Psychophysical Rating Results.

Rating indexes (bars) plotted as a function of the 6 stimulus conditions averaged over all participants ( $N=7$ ). Grey bars refer to stimulus conditions in which the probing region was grey; black bars to conditions in which the probing region was black. Insets above the bars are schematic illustrations of the stimuli. The white regions in the illustrations indicate the inducers of which the luminance levels were modulated (two insets on the left: two inducers; two insets in the middle: inner inducer; two insets on the right: outer inducer). The luminance of the probing region in the insets refers to the fixed luminance of the probing region during the rating trials, which then could be adapted by participants to reflect experienced induction. Error bars (SEM) and bar sizes for the black probing region were close to zero as subjects did not perceive brightness changes during inducer luminance modulations.

sign was used for in-phase settings. Figure 2 shows the mean rating indexes in the 6 conditions (error bars represent 1 SEM) obtained in seven subjects.

### Functional Imaging Experimental Design

Functional runs in induction and control condition (Figure 1A) lasted each 460s. Eccentricity localizer functional runs lasted 290s, and consisted of an alternation of 11s baseline and 11s localizer stimulation blocks (three blocks for each localizer). Four participants were shown three runs in induction condition, then three runs in control condition, and finally two localizer runs. The other participants started with control runs, followed by induction and localizer runs. A single session thus contained eight runs and anatomical

scan. In three participants, additional data were collected in separate sessions using stimulus patterns in which outer or inner inducer was removed and replaced with constant black ( $1.1\text{ cd/m}^2$ ).

## **Scanning Parameters**

Structural and functional MRI scanning was performed on a 3T Siemens Allegra head scanner using standard procedures and bird-cage head coil. Slice positioning was directed at calcarine sulcus and lower gyrus. A high-resolution T1-weighted three-dimensional (3D) anatomical scan was acquired for each participant (MP-RAGE; TR, 2,250 ms; matrix size,  $256 \times 256$ ; 192 slices; in-plane resolution,  $1\text{ mm}^2$ ). To measure blood oxygen level-dependent (BOLD) contrast, standard T2\*-weighted gradient-echo echo-planar imaging (EPI) was used to acquire 14 slices (TR/TE, 1,000/30 ms; flip angle  $90^\circ$ ; slice thickness, 2 mm; no slice gap; matrix size,  $128 \times 128$ ; in-plane resolution,  $2\text{ mm}^2$ ). Chosen parameters maximized temporal and spatial resolution, at the price of being able to cover only 28 mm from top to bottom of slice block. Resulting coverage is insufficient to cover ventral and dorsal representations of early visual cortical areas. We chose to align the top of the block of slices to the calcarine fissure, so that data collection typically was limited to the contralateral upper quadrant representation in ventral V1, V2 and V3.

## **Preprocessing and Analysis of MR Images**

Anatomical images were transformed to 3D standardized space (Talairach & Tournoux, 1988) with a resampled voxel size of  $1\text{ mm}^3$ . Functional measurements were co-registered to pre-standardized 3D anatomical space and preprocessed using BrainVoyager QX (Brain Innovation, The Netherlands). Of each functional run, the first six images were discarded. Functional images were transformed to 3D standardized space (Talairach & Tournoux, 1988) with a resampled isotropic voxel size of  $2\text{ mm}^3$ . Preprocessing steps of functional time series included slice time correction, 3D correction for head volume movements (sinc interpolation) and temporal filtering, which included linear trend removal and a high-pass filter of 3 cycles per time course (BrainVoyager QX, Brain Innovation, The Netherlands).

The spatially standardized anatomical images were segmented according to the grey and white matter boundaries, and the cortical sheet of each hemisphere was then tessellated into a 3D surface representation (Kriegeskorte and Goebel, 2001; Goebel et al., 2006). The cortical surfaces were then

### *3 Brightness Induction*

inflated in order to push the sulci outwards, and flattened to achieve a 2D representation of the entire cortical surface of each hemisphere. The time-series of the volume were resampled to the cortical surfaces and slightly spatially smoothed using a 2mm smoothing kernel based on nearest vertex neighbors. The timeseries on the surface were used for functional analysis.

Time series analysis was conducted using multiple regression analysis in the frame of the general linear model (GLM) for fMRI (Worsley & Friston, 1995). For induction and control conditions, the GLM consisted of a single predictor (luminance changes relative to grey level in fixation periods; effect coding, 0 = baseline) and a covariate (relaxation period; dummy coding) for each functional run. For the localizer experiment, the model comprised three predictors that modelled the sequence of blocks of eccentricity localizers (dummy coding). All predictors were convolved with a two-gamma function to correct for the hemodynamic response delay.

## **Retinotopic Mapping and Functional Alignment**

Retinotopic mapping (polar angle mapping) was carried out using standard procedures (Engel, Glover, & Wandell, 1997; Linden, Kallenbach, Heinecke, Singer, & Goebel, 1999; Sereno et al., 1995). Polar mapping was used to delineate borders between V1, V2, and V3, which were mapped on inflated cortical renderings of individual participants. The borders were used as guidelines to draw an equipolar line from central to peripheral representations within each visual area (Figure 2B) in each hemisphere. Because vertices in the inflated brain representation preserve anatomical distances (Goebel et al., 2006), matching an appropriately drawn equipolar line with outcomes of the localizer experiment permits localizing activity distributions (Beta coefficients resulting from linear regression analysis) as a function of anatomical distance in mms. We observed large differences in cortical magnification among participants, hemispheres, and visual areas, and, hence, the length of sampling differed from case to case. To average across hemispheres, we functionally aligned the anatomically selected cortical paths within each visual area across hemispheres into a common space by non-linear morphing, similar to previously published approaches [e.g., (Forshed, Schuppe-Koistinen, & Jacobsson, 2003; Listgarten, Neal, Roweis, & Emili, 2005)]. Alignment, analysis and plotting of results were performed using custom-written routines in Matlab (The Mathworks, Natick, MA, USA). By functionally aligning the cortical paths, constrained by localizer data, our data became spatially comparable over hemispheres and could be collapsed. As a result, averaged Beta coefficients are shown as a function of anatomi-

cal distance. The method is described in more detail elsewhere (Jans et al., submitted).

## Results

### Psychophysics

Participants verbally reported seeing the brightness illusion most strongly and consistently when the probing region was grey, and only in cases in which the outer inducer changed luminance, alone or in conjunction with the inner inducer. The psychophysical ratings obtained in a separate session in the scanner supported the participants' verbal reports (see Figure 2). Ratings differed significantly from the initial brightness of the grey probing region immediately after a change in luminance of both inducers ( $T[6] = -6.7$ ,  $P = 0.0005$ ) or immediately after a luminance change in only the outer inducer ( $T[6] = -4.2$ ,  $P = 0.0055$ ). In both cases, the rated changes in brightness were in the direction opposite to the luminance changes of the inducers. Participants did not see the brightness illusion in the grey probing region when luminance changed only in the inner inducer ( $T[6] = 1.9$ ,  $P = 0.11$ ). Participants perceived no brightness changes in the black probing region after luminance changes in any of the inducers ( $Ps > 0.23$ ). Interestingly, the ratings suggested that the brightness illusion in the grey probing region was stronger when luminance changed simultaneously in both inducers, compared to only the outer inducer. Analysis of the difference in ratings between the two conditions was significant ( $T[6] = -2.2$ ,  $P = 0.034$ , one-tailed). There was no difference in the strength of the induction effect when the data for inducer luminance increases and decreases were compared ( $Ps > 0.27$ ).

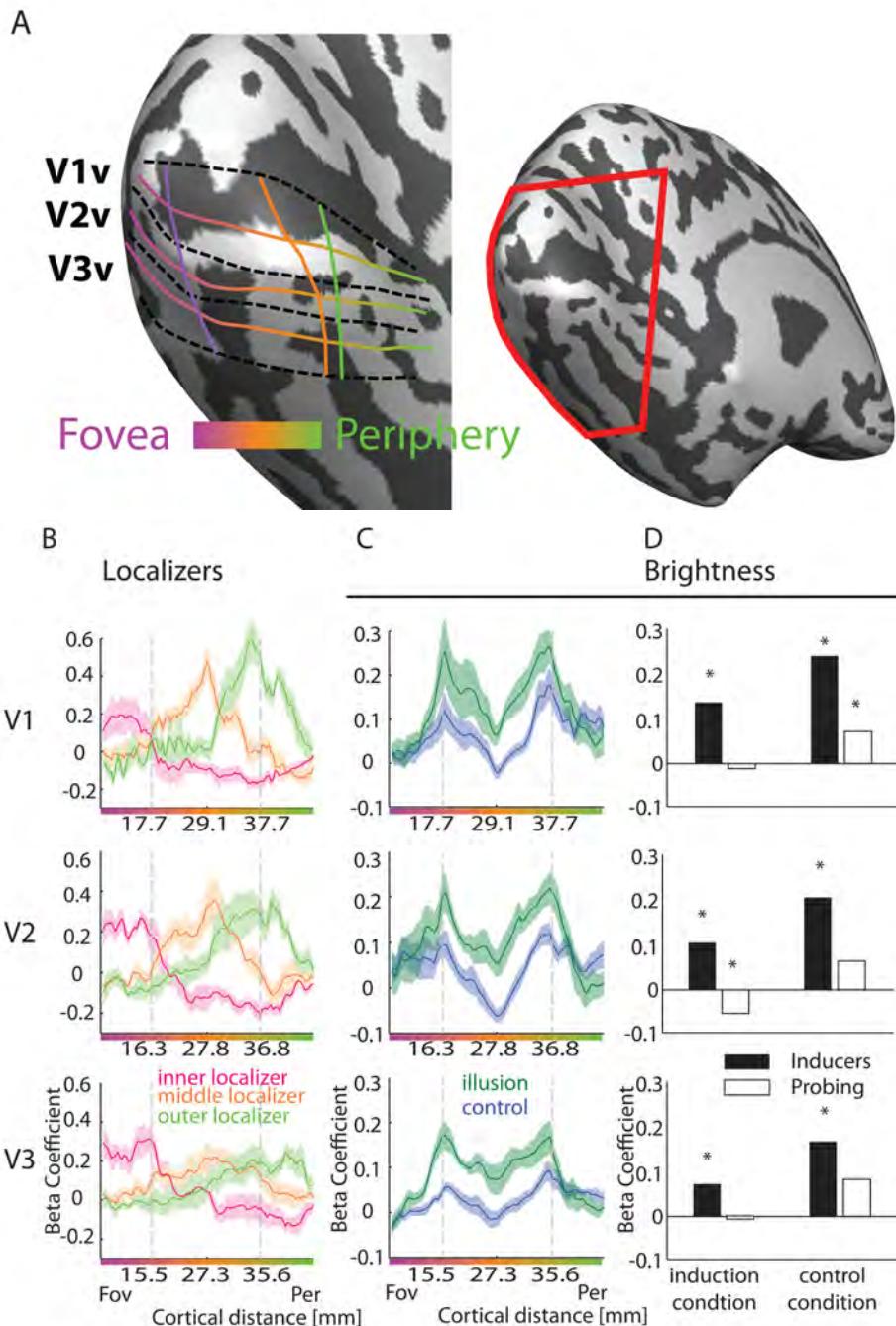
### Neural Correlates of Brightness Induction on the Cortical Surface

We analyzed fMRI signal during the brightness induction and control conditions sampled along cortical paths drawn in the ventral part of early visual areas V1, V2 and V3 to which data collection was limited (see Methods). In the remainder of the paper we use V1, V2 and V3 as shorthand for the ventral part of these visual areas. Data were sampled along cortical paths through the stimulus representations in V1, V2 and V3 (Figure 3A) in 14 hemispheres from 7 participants. After alignment, Beta coefficients were averaged along the 14 anatomical paths drawn through the stimulus

### *3 Brightness Induction*

representations in each area (see Methods). The resulting distributions of Beta coefficients show the correlation of fMRI signal with repeated localizer presentations or repeated inducer luminance modulations (Figure 3B-F).

Figs 3B and C show the Beta coefficients in localizer and brightness induction experiments sampled across the cortical paths in V1, V2 and V3. Beta coefficients for the inner (purple line), middle (orange) and outer localizers (green) are shown in Figure 3B. The peaks indicate the approximate anatomical location of the localizer stimuli, and distributions of Beta coefficients show a large spread away from the peaks of the localizer stimuli. Figure 3C shows the coefficients of the brightness induction (blue) and control condition (green) in the main experiment. We applied a repeated measures analysis of variance (ANOVA) to the data in Figure 3C, with within-subject factors Visual Area (V1, V2, V3), Condition (illusion, control) and Stimulus Region (probing region, inducers) to the Beta coefficients. We found significant main effects for Condition ( $F[1,6] = 66.6$ ,  $P < 0.001$ ) and Stimulus Region ( $F[1,6] = 36.3$ ,  $P = 0.001$ ), and a significant Visual Area x Condition interaction effect ( $F[2,12] = 4.5$ ,  $P = 0.035$ ). In the control condition (green line), in which participants reported no brightness induction (see Figure 1), mostly in-phase fMRI signal was found in the cortical projection areas of the localizers and probing region in all three visual areas (i.e., signal correlated positively with inducer luminance modulations; see Table 1). The in-phase signal was strongest for inducers and became weaker towards the middle of the probing region (Inducer > Probing; (V1:  $T[6] = 5.0$ ,  $P = 0.0025$ ; V2:  $T[6] = 2.9$ ,  $P = 0.0282$ ; V3:  $T[6] = 2.2$ ,  $P = 0.069$ ). At the probing region center, fMRI signal remained in-phase with the luminance changes (i.e., average beta coefficients were of positive sign) in all three visual areas (minimum  $T = 1.9$ ; see Table 1). In the induction condition (blue line in Figure 3), in which participants reported anti-phase brightness induction, beta coefficients at the probing region were significantly smaller than coefficients at the inducers in all three visual areas (V1:  $T[6] = 9.3$ ,  $P < 0.001$ ; V2:  $T[6] = 8.6$ ,  $P < 0.001$ ; V3:  $T[6] = 4.3$ ,  $P = 0.005$ ). In addition, however, three findings suggest the presence of another, anti-phase signal in the probing region associated with surface brightness perception. First, there was a strong decline of Beta values in the probing region during brightness induction compared to the control condition, as is evidenced by the divergence of green (control condition) and blue lines (induction condition) (Control > Induction at probing region; V1:  $T[6] = 4.8$ ,  $P = 0.003$ ; V2:  $T[6] = 4.6$ ,  $P = 0.004$ ; V3:  $T[6] = 2.5$ ,  $P = 0.045$ ). Second, this divergence was statistically significant not only in the probing region, but



### 3 Brightness Induction

	Induction				Control			
	Inducers		Probing Region		Inducers		Probing Region	
	T	P	T	P	T	P	T	P
V1	5.2	0.0021	-0.8	0.4292	4.8	0.0032	3.0	0.0232
V2	5.0	0.0026	-3.7	0.0101	5.6	0.0014	1.9	0.1023
V3	4.8	0.0029	-0.5	0.6179	5.9	0.0010	2.3	0.0642

**Table 1.** fMRI signal in inducers and probing region. Tests were performed on the Beta coefficients sampled along the cortical paths in visual areas V1, V2 and V3.

**Figure 3.** Cortical responses for the localizer and brightness stimuli.

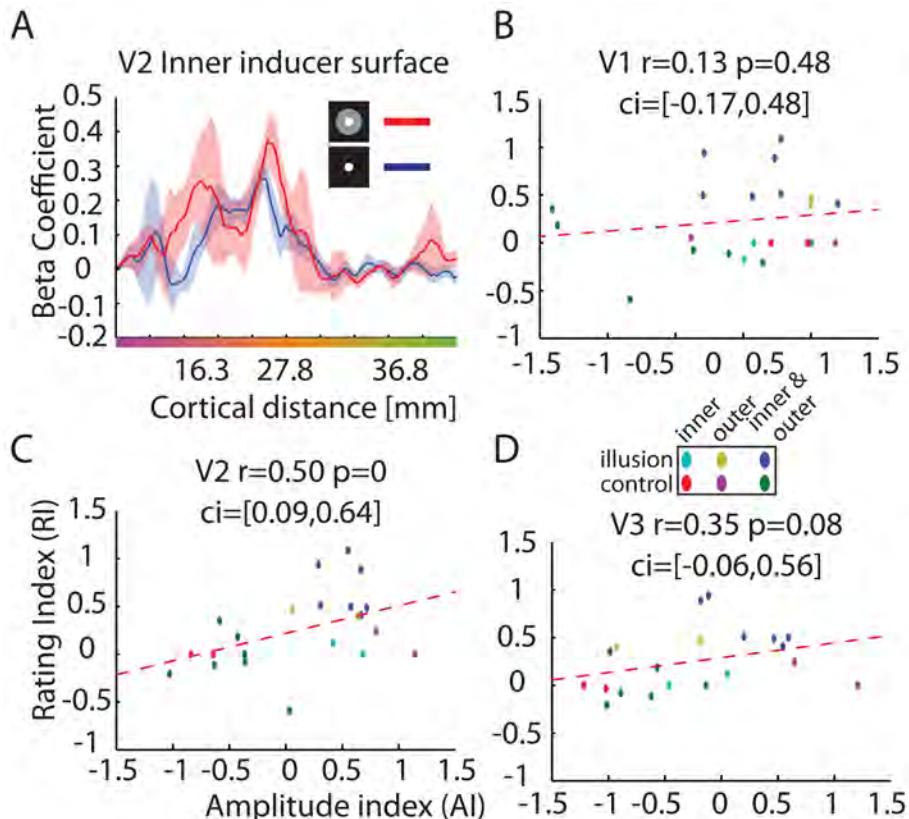
- (A) illustration of cortical paths along which functional data were sampled in V1, V2 and V3. The cortical paths contain a colour gradient (purple to green) corresponding to the continuum from fovea to periphery. (B-F) Overview of fMRI results for localizer and brightness induction experiments. Rows are plots for V1 (upper row), V2 (middle row) and V3 (bottom row).
- (B) Beta coefficients of the inner (purple), middle (orange) and outer (green) localizers sampled along the cortical paths (x-axis is marked with the same colour gradient as the cortical paths in figure 3A) plotted as a function of cortical distance (mm). Each localizer shows one peak of positive beta coefficients and a decline away from the peaks.
- (C) Beta coefficients of the brightness induction (blue) and control (green) stimuli sampled along the cortical paths plotted as a function of cortical distance (mm). Shaded regions represent  $\pm 1$  SEM ( $N=7$ , hemispheres averaged within participants). Conventions as in column (B).
- (D) Means of beta coefficients in inducers (black) and probing region (white).

also in the inducer region (Control > Induction at inducers; V1:  $T[6] = 2.7$ ,  $P = 0.035$ ; V2:  $T[6] = 4.4$ ,  $P = 0.004$ ; V3:  $T[6] = 2.9$ ,  $P = 0.029$ ). Third, the average of the coefficients was negative at the cortical center of the probing region for all three visual areas (all  $Ts < -0.5$ ), but was significant only for V2 ( $T[6] = -3.7$ ,  $P = 0.01$ ), but not for V1 ( $P = 0.4$ ) or V3 ( $P = 0.6$ ) (see Table 1). Taken together, this analysis indicates that in V2, but not in V1 or V3, there is an anti-phase signal in early visual cortex related to brightness induction in the probing region that due to spatial spread also attenuated in-phase activity of inducer regions.

## Control Experiments

The observation that in V2, anti-phase activity during induction was lost during the control condition strongly suggests that activity in early visual cortex contributed to the encoding of brightness in the probing region. However, local contrast changes at the border between inducers and probing region were different in the induction and control conditions (see Experimental Procedures and Figure 1A). Prior fMRI studies have shown that local changes in activity at one location in the retinotopic map can lead to opposite changes in activity in surrounding regions (Shmuel, Augath, Oeltermann, & Logothetis, 2006), possibly due to long-range lateral inhibition (Angelucci et al., 2002; Levitt & Lund, 2002). Hence, the specific local contrast changes in the induction condition may have led to anti-phase activity inside the probing region, while, for unknown reasons, the specific local contrast changes in the control condition may not have. Under this scenario, the anti-phase fMRI activity in the probing region during brightness induction might have been an artifact of local contrast conditions, rather than a correlate of surface brightness induction. To exclude this possibility, a stimulus was devised in which the outer inducer was replaced with constant black, and only the luminance of the inner inducer was manipulated. In this stimulus, there is no brightness induction neither for the grey nor for the black probing region (see rating results and Figure. 2), yet local contrast changes at the border between probing and inner inducer regions were identical to those in the original induction and control stimuli. If local contrast changes indeed determined the fMRI response irrespective of experienced surface brightness, then anti-phase activity would be expected for the grey, but not the black probing region, mimicking results with two inducers. However, if the perception of brightness induction were the determining factor for anti-phase activity in the probing region, then anti-phase activity would be expected neither for the grey nor the black probing region. Our results in V2 support the latter hypothesis, as the responses to the inducer next to the black and grey probing regions were indistinguishable (Figure 4A). This confirms that differences in local contrast changes between brightness induction and control stimuli are insufficient to explain the anti-phase activity observed during brightness induction (with a grey probing region and two inducers).

### 3 Brightness Induction



**Figure 4.** fMRI signal in brightness induction. (A) Beta coefficients of luminance modulations of the inner inducer-only condition shown as a function of V2 cortical distance. Responses to the stimulus with the grey (black) probing region are shown in red (blue) (shaded regions 1 SEM). Activity did not differ between conditions at any cortical position. (B-D) Rating index (RI) as a function of amplitude index (AI) from selected probing region voxels (see Figure 3) plotted for 3 stimulus conditions with the grey probing region in V1 (B), V2 (C) and V3 (D) during luminance up and down modulations. Colored symbols refer to participant data for induction conditions (see legend inset). Number of participants (N) for three types of brightness induction experiments: N(inner) = 3, N(outer) = 3, N(inner & outer) = 7.

## **Association Between Ratings and Neural Correlate of Brightness Illusion**

Finally, to investigate the perceptual relevance of anti-phase activity in the probing region, we tested whether the strength of anti-phase activity predicted the strength of brightness induction. We correlated the ratings of the participants and conditions for which we obtained functional imaging data. Results are shown in the scatterplots of Figure. 4B-D, with the correlation value, 95% confidence interval (CI) and associated P-value listed at the top of each panel. Ratings correlated significantly with the AI across the conditions only in area V2 ( $r = 0.5$ ,  $P < 0.001$ ), but not in V1 ( $r = 0.13$ ,  $P = 0.48$ ) or in V3 ( $r = 0.35$ ,  $P = 0.08$ ). Thus, fMRI activity in the probing region's representation in V2 predicted the perceived strength of induced surface brightness in the different stimulus conditions, but not in V1 and V3.

## **Discussion**

Using fMRI, we have investigated whether early visual cortex shows a functional neural correlate for brightness induction. We found an fMRI correlate of the anti-phase property of brightness induction in early human visual area V2. Furthermore, this correlate predicted the strength of experienced surface brightness modulations in a physically constant probing region. Our data converge with findings from single cell recording and hemodynamic measurement studies in animals that demonstrated a correlate of brightness induction in V2 (Roe studies), and suggest that brightness induction reveals a basic property of surface brightness perception, related to other types of contextual surface perception (Komatsu, 2006; Pessoa et al., 1998). Further, we demonstrated that our results are unlikely to reflect unspecific spread of activity generated by contrast changes at the border between inducers and probing regions. Instead our data suggest the existence of a small signal related to the perception of surfaces, in addition to a large signal related to the encoding of boundaries. Our results are unlikely to be associated with variations in pupil dilation as a result of luminance manipulations. Previous studies did not find differences in the retinotopic distribution of visual cortical activity after changes in pupil dilation (Haynes et al., 2004; Rooney & Cooper, 1988).

The finding of a correlate of brightness induction that was stronger in V2 than in V1 is in agreement with a number of other studies that have used other paradigms to manipulate brightness of neighboring surface of identi-

### *3 Brightness Induction*

cal luminance. For example, several brightness perception studies used a static brightness induction display (Craik-O'Brian-Cornsweet (COC) stimulus). Neuronal recording and optical imaging experiments in monkeys that were presented with a COC display demonstrated that the activity of V2 interstripe neurons reflected the illusory differences in COC-induced surface brightness, while V1 blob neurons did not (Roe et al., 2005). A recent human fMRI study confirmed that extrastriate cortex may contribute more to surface brightness perception in COC stimuli than V1 (Perna, Tosetti, Montanaro, & Morrone, 2005). Furthermore, neurophysiological studies of texture filling-in in a Troxler fading paradigm also demonstrated contributions of extrastriate but not striate areas (De Weerd et al., 1995). The stronger contribution of V2 compared to V1 also has been described in the domain of contour perception (Peterhans & von der Heydt, 1989, 1991; von der Heydt & Peterhans, 1989; von der Heydt, Peterhans, & Baumgartner, 1984). The evidence suggests that V2 contributes importantly to contextual processes that contribute to contour and surface perception.

However, several animal neurophysiological studies (Rossi et al., 1996; Rossi and Paradiso, 1999; Kinoshita and Komatsu, 2001) have shown a contribution of V1 to brightness induction. In addition, a few recent fMRI studies showed a brightness induction correlate in human V1 using a dynamic (Pereverzeva and Murray, 2008) as well as a static brightness induction stimulus (Boyaci et al., 2007). Our results do not exclude the possibility that V1 may also contribute to the perception of brightness, although this effect may be smaller compared to V2, and our paradigm or stimulus may not have been sensitive enough to pick up the effect in V1. Some fMRI studies in fact did not find a brightness correlate in either V1 nor extrastriate cortex using brightness induction (Cornelissen et al., 2006; Perna et al., 2005) as experimental paradigms.

Experimental design and its interaction with known limitations in temporal and spatial resolution (Bandettini, 2002; Kim, Richter, & Ugurbil, 1997; Menon & Kim, 1999) may be a factor that explains the different results in different studies of brightness induction. In the fMRI studies by Pereverzeva and Murray (2008) and Cornelissen et al. (2006), changes in inducer luminance were presented at a relatively high temporal rate of 1 Hz. While this presentation rate could be resolved in neurophysiological measurements, it may have reduced the sensitivity to pick up small surface-related signals in the presence of large boundary related signals (Boynton et al., 1996; Logothetis et al., 2001). The problem of lack of sensitivity might have been worsened by putative spatial or lateral spread of signals, which can be ex-

pected to be in counterphase. Nevertheless, it is difficult to determine why, despite very similar designs, the study from Pereverzeva and Murray (2008) and (Cornelissen et al., 2006; Pereverzeva & Murray, 2008) led to different conclusions.

Further, spatial resolution limitations of fMRI likely have determined in important ways with the fMRI correlate of brightness induction we have observed. Estimates of the spatial spread of the fMRI signal vary from 2 to 7mms HWHM (Cornelissen et al., 2006; Engel et al., 1997; Sereno et al., 1995). Our own data are rather in agreement with the estimates from Cornelissen et al. (2006). Several factors may contribute to the fMRI signal spread, including receptive field sizes in the cortical area under study (Dumoulin & Wandell, 2008; Smith, Singh, Williams, & Greenlee, 2001), neuronal signal spread due to lateral connectivity and feedback loops (Angelucci et al., 2002; Levitt & Lund, 2002) as well as hemodynamic factors (Buxton et al., 1998; Malonek & Grinvald, 1996). Because of this, fMRI signals generated in close spatial proximity will spread into each other's territory. Therefore, in our fMRI study anti-phase activity associated with brightness induction not only competes with in-phase activity related to luminance surfaces, but also with signals that encode the boundary of the probing region. This mix of signals related to physical aspects of the inducers is unrelated to the fMRI signal associated with brightness induction, and could contribute to the masking of the signal related to brightness induction. If this were true, then decreased fMRI point spread should be associated with increased magnitude of anti-phase activity in the probing region during brightness induction. We tested this hypothesis by computing in each participant the point spread of signals (in mms HWHM) from the inner and outer inducers into the probing region in the control condition, and comparing these with the Beta coefficients in the probing region during brightness induction. We found that larger brightness induction responses (i.e., negative beta coefficients further away from 0) correlated with smaller signal spread from the inducers into the cortical representation of the probing region ( $r=0.59$ ,  $P=0.042$ ). Another sign of the interaction between spread and induction is the finding that in-phase responses to inducer luminance modulations overall were smaller in the induction condition than in the control condition. This may be due to antiphase activity present in the probing region during induction spreading into the territory of inducers thereby attenuating in phase activity from inducers.

The precise neural circuitry that underlies brightness induction is currently unknown. Brightness induction may result from contextual interactions

### *3 Brightness Induction*

among neurons that are sensitive to physically homogenous stimulation, localized in V1 blobs and V2 thin stripes, under the assumption that these interactions only take place among surfaces that sufficiently drive stimulated neurons (i.e. not when one of the surfaces is black). Alternatively, or in addition, contrast-polarity sensitive neurons (Zhou, Friedman, & von der Heydt, 2000) may contribute to induction, or to filling-in effects in which a signal at boundaries is interpolated into surfaces (Grossberg, 2003). Furthermore, it is possible that linear filtering explains a significant portion or all of the induction effects. Blakeslee and McCourt (2004) have shown that many brightness effects can be explained by assuming multi-scale spatial filtering in the human visual system combined with contrast normalization. In this proposal, neurons in V1 blobs and V2 thin stripes might play a special role in low spatial frequency filtering, rather than being a substrate for spreading of surface information (Goebel & De Weerd, 2009; Grossberg, 1987a, 1987b, 2003). This is related to Cornelissen et al.'s suggestion that surface-related information is derived from spatial filtering, without the need for spreading activity. However, our fMRI data do not permit to distinguish between the various possible mechanisms.

Although participants were required to carefully fixate, which by itself may be an attention-demanding task (Martinez-Conde, Macknik, Troncoso, & Dyar, 2006), the possibility exists that at least some attention was covertly allocated to the peripherally presented stimuli, which might have influenced the described effects. However, neural signals related to surface perception have been observed in prior studies in the absence of attention (Boyaci, Fang, Murray, & Kersten, 2007; Meng, Remus, & Tong, 2005; Weil, Watkins, & Rees, 2008), and even under anesthesia (Roe et al., 2005; Rossi & Paradiso, 1999; Rossi et al., 1996). Therefore, we consider it unlikely that the reported data reflect predominantly attention effects.

Research into the neural correlates of surface perception has yielded conflicting data and views. For example, a neurophysiological study on surface texture in monkey visual cortex provided evidence for surface-related signals in V2 and V3 (De Weerd, Gattass, Desimone, & Ungerleider, 1995), but a recent fMRI study did not confirm that evidence (Weil et al., 2008). An fMRI study on color filling-in reported a correlate in human V1 (Sasaki & Watanabe, 2004), but a neurophysiological study in monkey V1 and V2 did not (Friedman, Zhou, & von der Heydt, 2003). Similar differences between studies remain in the domain of surface brightness perception, and further research is required to resolve them. The present study, however, supports the idea that there is a signal related to surface brightness in extrastriate

cortex, although much more work is required to understand the mechanisms underlying that signal.

## References

- Angelucci, A., Levitt, J. B., Walton, E. J., Hupe, J. M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *J Neurosci*, 22(19), 8633-8646.
- Bandettini, P. A. (2002). fMRI: The Spatial, Temporal, and Interpretive Limits of Functional MRI. In K. L. Davis, D. Charney, J. T. Coyle & C. Nemeroff (Eds.), *Neuropsychopharmacology: The Fifth Generation of Progress: An Official Publication of the American College of Neuropsychopharmacology* (pp. 343-356). Philadelphia: Lippincott Williams & Wilkins.
- Boyaci, H., Fang, F., Murray, S. O., & Kersten, D. (2007). Responses to lightness variations in early human visual cortex. *Curr Biol*, 17(11), 989-993.
- Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci*, 16(13), 4207-4221.
- Buxton, R. B., Wong, E. C., & Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magn Reson Med*, 39(6), 855-864.
- 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. *J Neurosci*, 26(14), 3634-3641.
- De Weerd, P., Gattass, R., Desimone, R., & Ungerleider, L. G. (1995). Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature*, 377(6551), 731-734.
- Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage*, 39(2), 647-660.
- Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex*, 7(2), 181-192.
- Forshed, J., Schuppe-Koistinen, I., & Jacobsson, S. P. (2003). Peak alignment of NMR signals by means of a genetic algorithm. *Analytica Chimica Acta*, 487(2), 189-199.
- Friedman, H. S., Zhou, H., & von der Heydt, R. (2003). The coding of uniform colour figures in monkey visual cortex. *J Physiol*, 548(Pt 2), 593-613.
- Goebel, R., & De Weerd, P. (2009). Perceptual Filling-in: From Experimental Data to Neural Network Modeling. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences*. Cambridge, MS.: The MIT Press.
- Grossberg, S. (1987a). Cortical dynamics of three-dimensional form, color, and

### *3 Brightness Induction*

- brightness perception: I. Monocular theory. *Percept Psychophys*, 41(2), 87-116.
- Grossberg, S. (1987b). Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory. *Percept Psychophys*, 41(2), 117-158.
- Grossberg, S. (2003). Filling-in the forms: Surface and boundary interaction in visual cortex. In L. Pessoa & P. De Weerd (Eds.), *Filling-in: From Perceptual Completion to Skill Learning* (1st Edition ed., pp. 13-37). Oxford: University Press.
- Haynes, J. D., Lotto, R. B., & Rees, G. (2004). Responses of human visual cortex to uniform surfaces. *Proc Natl Acad Sci U S A*, 101(12), 4286-4291.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, 160, 106-154.
- Kim, S. G., Richter, W., & Ugurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magn Reson Med*, 37(4), 631-636.
- Kinoshita, M., & Komatsu, H. (2001). Neural representation of the luminance and brightness of a uniform surface in the macaque primary visual cortex. *J Neurophysiol*, 86(5), 2559-2570.
- Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nat Rev Neurosci*, 7(3), 220-231.
- Levitt, J. B., & Lund, J. S. (2002). The spatial extent over which neurons in macaque striate cortex pool visual signals. *Vis Neurosci*, 19(4), 439-452.
- Linden, D. E., Kallenbach, U., Heinecke, A., Singer, W., & Goebel, R. (1999). The myth of upright vision. A psychophysical and functional imaging study of adaptation to inverting spectacles. *Perception*, 28(4), 469-481.
- Listgarten, J., Neal, R. M., Roweis, S. T., & Emili, A. (2005). Multiple alignment of continuous time series. *Advances in Neural Information Processing Systems*, 17.
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843), 150-157.
- Malonek, D., & Grinvald, A. (1996). Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: implications for functional brain mapping. *Science*, 272(5261), 551-554.
- Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Dyar, T. A. (2006). Microsaccades counteract visual fading during fixation. *Neuron*, 49(2), 297-305.
- Meng, M., Remus, D. A., & Tong, F. (2005). Filling-in of visual phantoms in the human brain. *Nat Neurosci*, 8(9), 1248-1254.
- Menon, R. S., & Kim, S. G. (1999). Spatial and temporal limits in cognitive neuroimaging with fMRI. *Trends Cogn Sci*, 3(6), 207-216.
- Pereverzeva, M., & Murray, S. O. (2008). Neural activity in human V1 correlates with dynamic lightness induction. *J Vis*, 8(15), 8 1-10.
- Perna, A., Tosetti, M., Montanaro, D., & Morrone, M. C. (2005). Neuronal mechanisms

## References

- for illusory brightness perception in humans. *Neuron*, 47(5), 645-651.
- Pessoa, L., Thompson, E., & Noe, A. (1998). Finding out about filling-in: a guide to perceptual completion for visual science and the philosophy of perception. *Behav Brain Sci*, 21(6), 723-748; discussion 748-802.
- Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *J Neurosci*, 9(5), 1749-1763.
- Peterhans, E., & von der Heydt, R. (1991). Subjective contours—bridging the gap between psychophysics and physiology. *Trends Neurosci*, 14(3), 112-119.
- Roe, A. W., Lu, H. D., & Hung, C. P. (2005). Cortical processing of a brightness illusion. *Proc Natl Acad Sci U S A*, 102(10), 3869-3874.
- Rooney, B. J., & Cooper, R. M. (1988). 2-Deoxyglucose uptake in the rat visual system following pupil dilation with atropine. *Vision Res*, 28(2), 259-262.
- Rossi, A. F., & Paradiso, M. A. (1999). Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *J Neurosci*, 19(14), 6145-6156.
- Rossi, A. F., Rittenhouse, C. D., & Paradiso, M. A. (1996). The representation of brightness in primary visual cortex. *Science*, 273(5278), 1104-1107.
- Sasaki, Y., & Watanabe, T. (2004). The primary visual cortex fills in color. *Proc Natl Acad Sci U S A*, 101(52), 18251-18256.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268(5212), 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. *Nat Neurosci*, 9(4), 569-577.
- Smith, A. T., Singh, K. D., Williams, A. L., & Greenlee, M. W. (2001). Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cereb Cortex*, 11(12), 1182-1190.
- Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the humran brain. New York: Thieme Medical.
- von der Heydt, R., & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *J Neurosci*, 9(5), 1731-1748.
- von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, 224(4654), 1260-1262.
- Weil, R. S., Watkins, S., & Rees, G. (2008). Neural correlates of perceptual completion of an artificial scotoma in human visual cortex measured using functional MRI. *Neuroimage*, 42(4), 1519-1528.
- Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited—again. *Neuroimage*, 2(3), 173-181.
- Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *J Neurosci*, 20(17), 6594-6611.



## **4 Induced brightness filling-in in early visual areas: neural modeling of fMRI data in a common brain space**

based on: Peters, J.C., Jans, B., Van de Ven,V., De Weerd, P., and Goebel, R., (2010), Dynamic brightness induction in V1: Analyzing simulated and empirically acquired fmri data in a "common brain space" framework. *Neuroimage*, 52(3):973-984.

## Abstract

Computational neuromodeling may help to further our understanding of how empirical neuroimaging findings are generated by underlying neural mechanisms. Here, we used a simple computational model that simulates early visual processing of brightness changes in a dynamic, illusory display. The model accurately predicted illusory brightness changes in a grey area of constant luminance induced by (and in anti-phase to) luminance changes in its surroundings. Moreover, we were able to directly compare these predictions with recently observed fMRI results on the same brightness illusion by projecting predicted activity from our model onto empirically investigated brain regions. This new approach in which generated network activity and measured neuroimaging data are interfaced in a common representational “brain space” can contribute to the integration of computational and experimental neuroscience.

## Introduction

How surface perception is achieved is not well understood, and a lot of empirical and computational modeling work is devoted to tackle questions on its underlying neuronal mechanisms. The present paper focuses on the perception of surface brightness, and its aims are twofold. First, we aim to build a simple model that generates neuronal activity that correlates with the perception of surface brightness. The model generates specific predictions in the form of a dynamically changing topographic map of activity that captures specific aspects of surface brightness perception. Second, we aim to integrate empirical and computational research by representing empirical and modeled data in a common space. The approach we propose is to project the predicted activity in a modeled area onto corresponding cortical regions where the data will be collected, to achieve direct and quantitative hypothesis testing. This requires large scale neural network modeling combined with a strategy to link modeled units (cortical columns) with specific neuroimaging units (voxels, vertices) in the empirically obtained brain model. To achieve this, modeled and empirical units are connected in a common anatomical space. Here, we show the feasibility of that approach.

To demonstrate our modeling approach, we will focus on the paradigm of brightness induction, in which an intermediate luminance patch on a bright background appears darker than on a dark background (White, 1979). This illusion is maximized in a dynamic display developed by DeValois (De Valois

et al., 1986), in which a grey area of constant luminance is flanked by areas whose luminance varies sinusoidally between black and white. We refer to the gray area of constant luminance as the “probing region” as we use it to probe mechanisms of surface perception, and refer to the areas whose luminance is modulated as the “inducers” (see Figure 2). The brightness percept in the gray area with fixed luminance varies in anti-phase to sinusoidal changes in physical luminance of the inducers. The neuronal mechanisms that cause these illusory brightness changes are not well-known. The most convincing evidence comes from neurophysiological recordings in the cat (Rossi and Paradiso, 1999; Rossi et al., 1996), which showed that in about 10% of area 17 neurons with their receptive fields (RFs) placed within the probing region, activity was modulated in anti-phase to inducer luminance (i.e., activity correlated to perceived brightness changes). Two human studies using functional magnetic resonance imaging (fMRI) have reported data that appear in agreement with the neurophysiological data (Pereverzeva and Murray, 2008; van de Ven et al., submitted), whereas a third fMRI study (Cornelissen et al., 2006) did not observe a brightness induction correlate in early visual cortex.

The model we propose contains three main parts (Figure 1). It consists of an input layer (which in essence captures physical luminance values in the image), and gives input to a layer where contours are extracted from this input, and to another layer where measurements across surfaces are modulated by information spreading from contours. In broad terms, this approach is similar to the models of Grossberg and coworkers (Grossberg, 1987a, 1987b; 2003; Grossberg and Todorovic, 1988; Grossberg and Raizada, 2000; Raizada and Grossberg, 2001; see also Gerrits et al., 1966; Gerrits and Vendrik, 1970;) in which an interaction between contour representations (in the Boundary Contour System or BCS) and spreading of surface features (in the Feature Contour System or FCS) leads to the experience of ‘visible’ surfaces (e.g., Grimson, 1982;). The idea that there are separate but interacting streams of visual processing related to the processing of boundaries and the processing of surfaces is also supported by anatomical and neurophysiological evidence (see Discussion). For example, V1 blob regions contain neurons that are often well stimulated by homogeneous surfaces that extend beyond their RFs, while V1 interblob regions contain neurons that are preferentially stimulated by oriented contrasts (e.g. contours and edges). Without claiming that our model captures the complexity of early visual processing, we will refer to the input layer as the Model-retina ( $M-<\text{retina}>$ ), and the two other layers as the Model V1

#### 4 Brightness Induction Modeling

blobs ( $M-), and Model V1 interblobs ( $M-). Processing in our model thus occurs in two stages: In the first stage of processing, contour information is extracted from local discontinuities in visual input (in  $M-); in the second stage, in  $M-, contour-related activity from  $M- is integrated with  $M- activity reflecting surface-related input. More specifically, contours are used as seeding points from which modulatory activity spreads, affecting the luminance representation in line with the percept. Specific aspects of the present modeling approach that differ from earlier modeling efforts, and their physiological interpretation will be treated in the Discussion.$$$$$$

In parallel with the development of the computational model, we recently acquired fMRI data in an experiment in which induced brightness perception was correlated with activity in early visual areas (van de Ven et al., submitted). We used a stimulus design adapted from neurophysiological studies (Rossi and Paradiso, 1999; Rossi et al., 1996), in which the brightness perception in a probing region of constant luminance is modulated by surrounding inducer regions that dynamically vary in luminance. In the illusion condition, participants reported a strong brightness illusion in a grey probing region, which was in anti-phase with luminance modulations in the inducers. Participants did not report the illusion in the control condition in which the probe was black instead of grey. This induced brightness percept correlated with activity in retinotopically mapped clusters in striate and extrastriate visual cortex. That is, activity in the probing region changed in anti-phase to physical luminance modulations in inducers in the illusion condition, whereas no activity was observed when the probe surface was black. We investigated the extent to which our simple model predictions captured aspects of cortical activity in response to induced surface brightness.

To relate the output of our model directly to acquired fMRI data, we followed the “Common Brain Space” (CBS) framework (Goebel and Horwitz, in preparation, Goebel and De Weerd, 2009). In this new framework, the processing units of a neural network model are connected to points (vertices) of brain models (cortex meshes) implementing specific spatial hypotheses. As a result, a running network simulation creates time course data directly on a linked brain model allowing to relate predicted and measured topographic neuroimaging data in a highly specific manner, i.e. by using the same data analysis tools (e.g., general linear model or multi-voxel pattern analyses).

## Methods

### Brightness illusion and brightness control design

We presented a comparable stimulus design to the neural model and to human subjects (van de Ven et al., submitted). In the fMRI experiment, the stimulus (see example in lower left of figure 3) consisted of an annulus of constant luminance ( $6^\circ$  width; probing region) interspersed between a central disk ( $3^\circ$  width) and a peripheral annulus ( $6^\circ$  width; the inducers). The inducers of the stimulus changed luminance quasi-linearly in 2s (range =  $1.1 - 164 \text{ cd/m}^2$ ), whereas the probe region had a fixed grey ( $7.9 \text{ cd/m}^2$ ; induction condition) luminance. After each "up" (inducer change from black to white) and "down" (inducer change from white to black) event, the inducers stayed at maximal and minimal luminance respectively, for 5s. The control condition was identical to the brightness induction condition, except that the probe region was black ( $1.1 \text{ cd/m}^2$ ) instead of grey.

The experimental design was similar for the simulation experiments, except that the stimulus covered the entire "visual field", and luminance of the inducers changed quasi-linearly from 0.00-1.00 input with probe region luminance fixed on 0.50 (induction condition) or 0.00 (control condition).

### Large-scale neural network model

Our large-scale neural network model (Figure 1) is composed of three retinotopically organized, square layers representing the retina ( $M\text{-}<\text{retina}>$ ), and the blob ( $M\text{-}<\text{V1}_{\text{blob}}>$ ) and interblob ( $M\text{-}<\text{V1}_{\text{inter}}>$ ) areas in V1. In the employed CBS framework, individual processing units are not representing single neurons but are conceptualized as "cortical column units" (CCUs). We believe this is the appropriate level of modeling, as measuring at the level of cortical columns - the building blocks of the visual system (Hubel and Wiesel, 1959, 1962) - is becoming in reach when employing state-of-the-art ultra high-field, fMRI measurements (Yacoub et al., 2008). The CCU's dynamics are defined by the subsequent equations (Goebel, 1993):

$$\begin{aligned} \text{net}_i(t) &= \sum_{j=1}^n w_{ij} a_j \\ a_i(t) &= (1 - \tau)a_i(t - 1) + \tau\sigma(\text{net}_i(t) + b_i) \end{aligned}$$

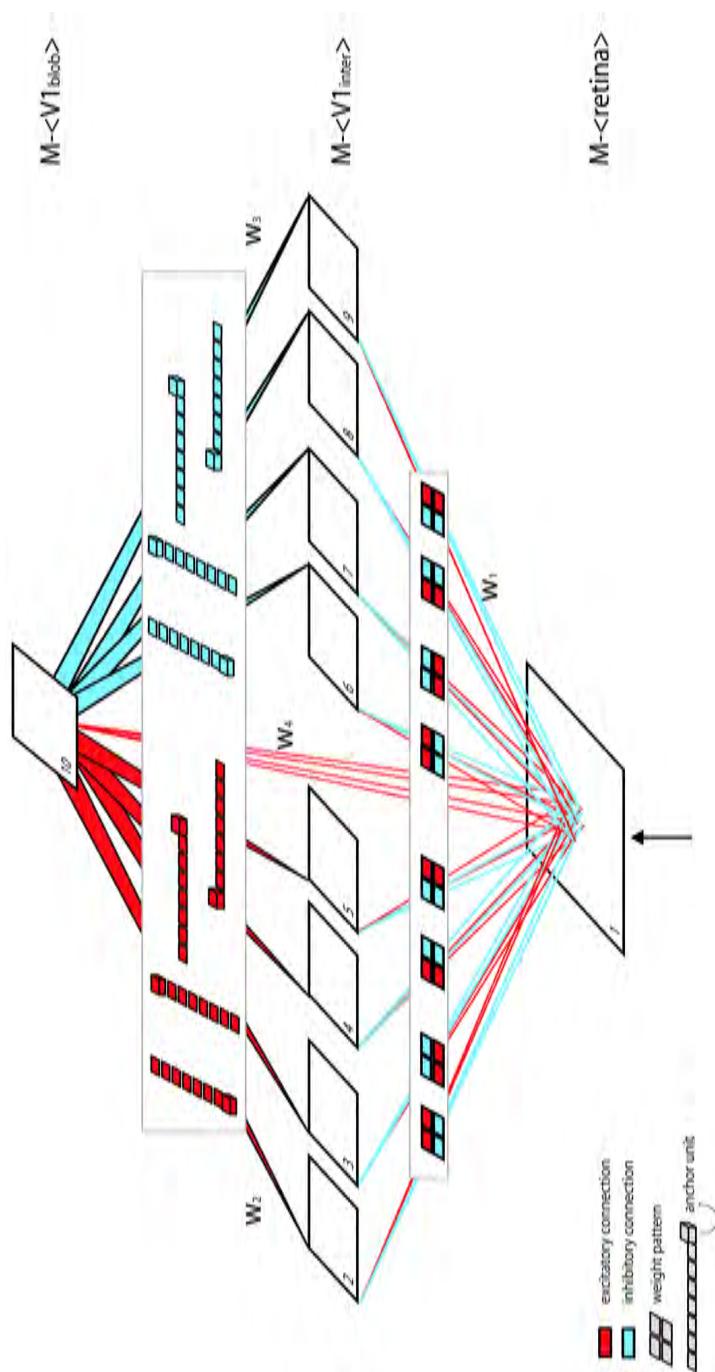
where  $a_i(t)$  is the average spike output of unit  $i$  at time  $t$ ,  $\text{net}_i(t)$  is the net input (excitatory minus inhibitory input) for unit  $i$  at time  $t$ ,  $b_i$  is a

## 4 Brightness Induction Modeling

bias term, and  $\sigma(x)$  is the logistic (sigmoidal) function. Finally, the value  $w_{ij}$  is the directed weight from CCU  $j$  to CCU  $i$ . The value  $\tau(0 < \tau \leq 1)$  determines how strongly the activation value (average spiking activity) at the last time point ( $t-1$ ) is influencing the activity of a unit at the current time point  $t$ .

Area M-<retina> contains 256x256 CCUs representing retinotopically organized cortical columns signaling physical surface luminance. In M-<V1>, the number of CCUs is a factor of four lower as in M-<retina> (Burt and Adelson, 1983), corresponding to an increased receptive field size. The M-< $V1_{inter}$ > area is subdivided in 8 layers with each layer representing cells with a specific “contour detection sensitivity” (i.e., a specific combination of orientation sensitivity and contrast polarity). As depicted in Figure 1, each of the cells in the M-< $V1_{inter}$ > area receives two inhibitory and two excitatory inputs from M-<retina>, the combination of which creates a receptive field with a specific contour detection sensitivity. Specifically, cells in M-< $V1_{inter}$ > that receive two excitatory inputs on the left, right, bottom, or upper side of the 2x2 array of inputs, correspond to respectively left, right, bottom or top contour detection cells. Although these cells are intermingled within the interblob area of V1, they are displayed in separate layers for clarity. In addition to the input from M-<retina>, each CCU receives (a weak) inhibitory input from its eight direct neighbors within the same layer. This lateral inhibition is modeled as a Gaussian distribution with a weight of -0.3 and a decay of 0.5. Finally, area M-< $V1_{blob}$ > integrates the contour information of M-< $V1_{inter}$ > with the brightness information signaled by area M-<retina>. Each CCU of this area receives four excitatory connections from M-<retina>, which is modulated by inhibitory or excitatory input from M-< $V1_{inter}$ >, as depicted in Figure 1. Furthermore, a lateral spreading (with weight 0.3 and decay 0.5) was implemented in M-< $V1_{blob}$ >. This connection scheme leads to edge-induced brightness filling-in in area M-< $V1_{blob}$ >. Thus, whereas area M-<retina> represents physical luminance information, neural activity in area M-< $V1_{blob}$ > should be more akin to the perceived brightness of visual input. The connection weights from M-<retina> to M-< $V1_{inter}$ > and from M-<retina> to M-< $V1_{blob}$ > were 4, and 0.5, respectively. In addition, the first and second halves of the M-< $V1_{inter}$ > layers were connected to M-< $V1_{blob}$ > with a weight of 4 and -4, respectively.

To increase the comparability between the two data sets in the temporal dimension, we transformed the CCU activity from spiking to BOLD signals by applying the following computations on the network activity: First we



## 4 Brightness Induction Modeling

**Figure 1.** Model Architecture. The model is composed of three areas. Physical luminance information enters the first area, M-<retina>, and is directly conveyed to the third area, M-<V1<sub>blob</sub>>, via connection weight w4. Through the group of w1 weights, the activity of four M-<retina> CCUs converges on one CCU of the second area, M-<V1<sub>inter</sub>>. CCUs in this area are clustered in eight sheets (2-9) according to their “contour detection” sensitivity, as indicated by the connection pattern of w1 (blue and red rectangles) M-<V1<sub>inter</sub>> is connected to M-<V1<sub>blob</sub>> with excitatory (w2) or inhibitory (w3) connections depending on the contour detection sensitivity of the M-<V1<sub>inter</sub>> cluster. Note that the rows of colored CCUs indicating the connection patterns of these weights represent 40 CCUs. In M-<V1<sub>blob</sub>> contour and luminance information is integrated, resulting in activity reflecting perceived brightness. In layer two to nine, there is local lateral inhibition (not shown) with strength decaying over distance following a Gaussian distribution. In layer ten, there is a lateral Gaussian excitation.

calculated the integrated synaptic activity,  $net_i^{abs}$ , as the sum of the absolute activity arriving via excitatory and inhibitory synapses at CCU  $i$  (which differentiates  $net_i^{abs}$  from the standard ‘netto’ input  $net$ ):

$$net_i^{abs} = |\sum w_{ij}^{ex} a_j| + |\sum w_{ij}^{inh} a_j|$$

Subsequently, we computed the proto-fMRI signal by integrating synaptic and spike activity (akin to the local-field potentials, LFP):

$$a_i^{fMRI} = \lambda net_i^{abs} + (1 - \lambda) a_i$$

with the weight factor  $\lambda$  set to 0.8, as LFP and fMRI data have been shown to predominantly reflect (subthreshold) synaptic activity (Logothetis et al., 2001; Maier et al., 2008). Then, we transformed this signal into the fMRI signal  $y_i$  by convolving it with a two-gamma hemodynamic response function HRF<sub>i</sub> (Friston et al., 1998):

$$y_i = HRF_i(a_i^{fMRI})$$

We finally mimicked physiological, measurement, and instrumental noise in the BOLD response by incrementally adding 30% of white Gaussian noise ( $N(t)$ ) (Roebroeck Formisano, and Goebel, 2005). Note that for any time instants  $t_i$ , the corresponding  $N(t_i)$  values are independent Gaussian random variables:

$$y_{noise} = y_i + N(t) \text{ with } E[N(t)] = 0$$

The network performed ten runs (460 volumes each) with the brightness induction design and one run with the brightness control design (290 volumes each) similarly to the human fMRI study described below.

## fMRI data

Seven healthy adults participated in the fMRI brightness induction, control, and (stimulus) localizer experiments. In addition to six functional runs for the brightness induction and control conditions (460 volumes each), two functional runs (290 volumes each) were acquired to localize the retinotopic representation of the inducer and probe regions using standard retinotopic mapping procedures.

Standard T2\*-weighted gradient-echo echoplanar imaging were acquired (3T Siemens Allegra; TR/TE: 1,000/30 ms; flip angle: 90°; slice thickness: 2 mm; 128×128 matrix) and submitted to standard preprocessing steps using BrainVoyager QX (Brain Innovation, The Netherlands). Based on the independent stimulus localizer Patches-of-Interest (POIs) were defined for each subject, constituting the retinotopic representations of the probing region and abutting inducer stimulus surfaces in area V1 and V2. Conditional effects between V1 POIs timecourses were analyzed in a similar fashion as the simulated M-<V1<sub>blob</sub>> activity as described in the next section. In addition, a high-resolution anatomical scan (T1; MPRAGE) was measured for each participant, and a cortical mesh (constituting the white-gray matter boundary of the cortical sheet) was reconstructed, inflated and flattened (see Figure 6 for an example). Additional details are described in van de Ven et al. (submitted).

## fMRI and neural net integration in Common Brain Space framework

Figure 2 shows the essential features of the CBS approach. To enable the direct comparison between our model's predictions and experimental fMRI

#### *4 Brightness Induction Modeling*

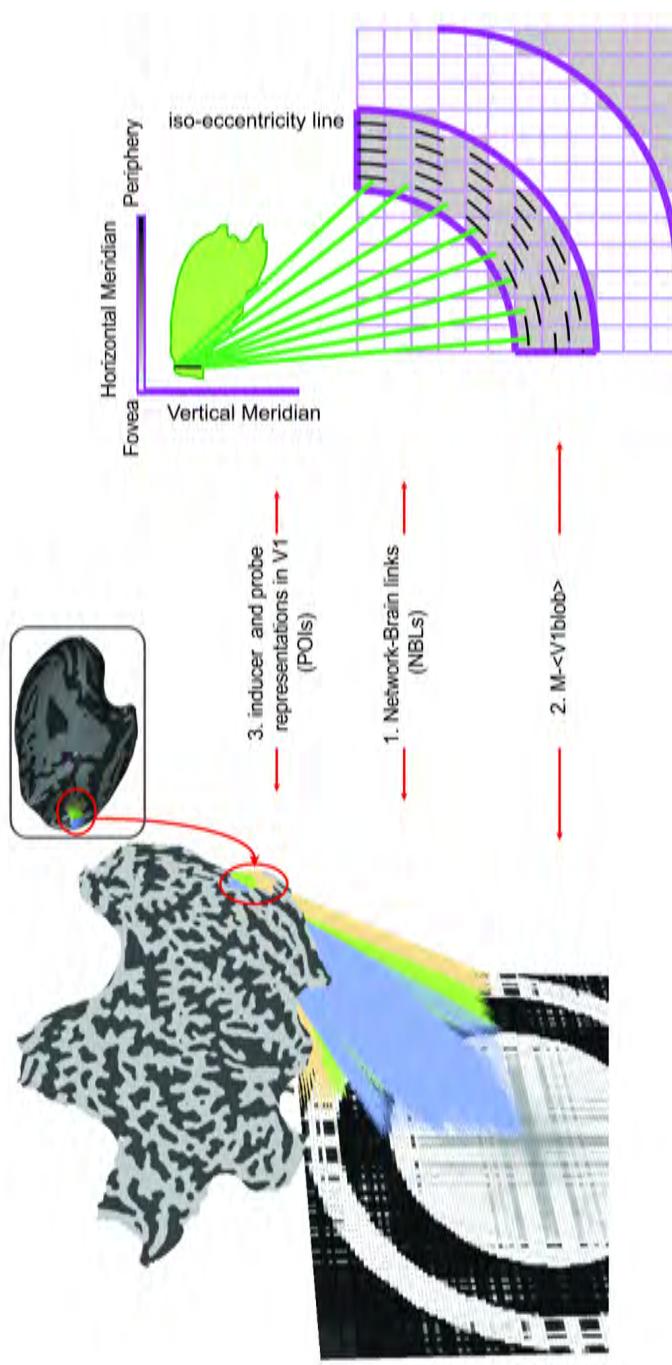
data, we projected our model on a cortical mesh of a subject who participated in our fMRI study (left side of Figure 2). On this mesh, we marked the subject-specific V1 POIs (brown, green, and blue patches in Figure 2). Then, we connected each CCU with the topographically corresponding patch in the visual cortex via a Network-Brain link (NBL; see right side of Figure 2). Via these NBLs, CCU activity can be directly projected on the cortical sheet at each time step of a running simulation.

Thereby, complete time series of network activity can be created with a similar data format as measured fMRI time series. This enables the analyses of activity changes in simulated and measured data using identical analyses methods. To illustrate this, we created time series of V1 POI activity (Figure 6) as provided by CCU ynoise activity in the upper quadrant of  $M- (Figure 2), plotted event-related averages (Figure 7), run a General Linear Model (GLM; Figure 8), and compared results based on the observed V1 and simulated  $M- data (Figure 9). The GLM design matrix contained a single predictor to model luminance changes (relative to grey level during rest periods) and was convolved with a two-gamma hemodynamic impulse response function to account for the hemodynamic delay.$$

All modeling was performed with Neurolator 3D, a neural network simulation software package, which can be conveniently interfaced with fMRI data analyses software such as BrainVoyager QX (Brain Innovation BV, Maastricht, The Netherlands).

## Results

The current model provides a simulation of brightness processing in the primary (V1) visual cortex. The network's connection scheme is explained in Figure 3, in which the model areas display schematic activation profiles in response to a stimulus with inducers that are brighter than the probing region. Depicted grey levels indicate the strength of CCU activity. Several CCUs are enlarged and shown on top of the activity profiles to illustrate the activity flow. The first layer,  $M-, represents physical luminance information. This information is processed in two subsequent stages: In the first stage, contours are extracted from local luminance discontinuities, which are signaled from  $M- to  $M-. Activity in  $M- depends on the contour detection sensitivity, specified by the connection pattern (see colored figures accompanying w1 weights). An  $M- CCU connected to four  $M- CCUs that represent a homogeneous surface will receive a balanced input of excitation and inhibition. In contrast,$$$$$$

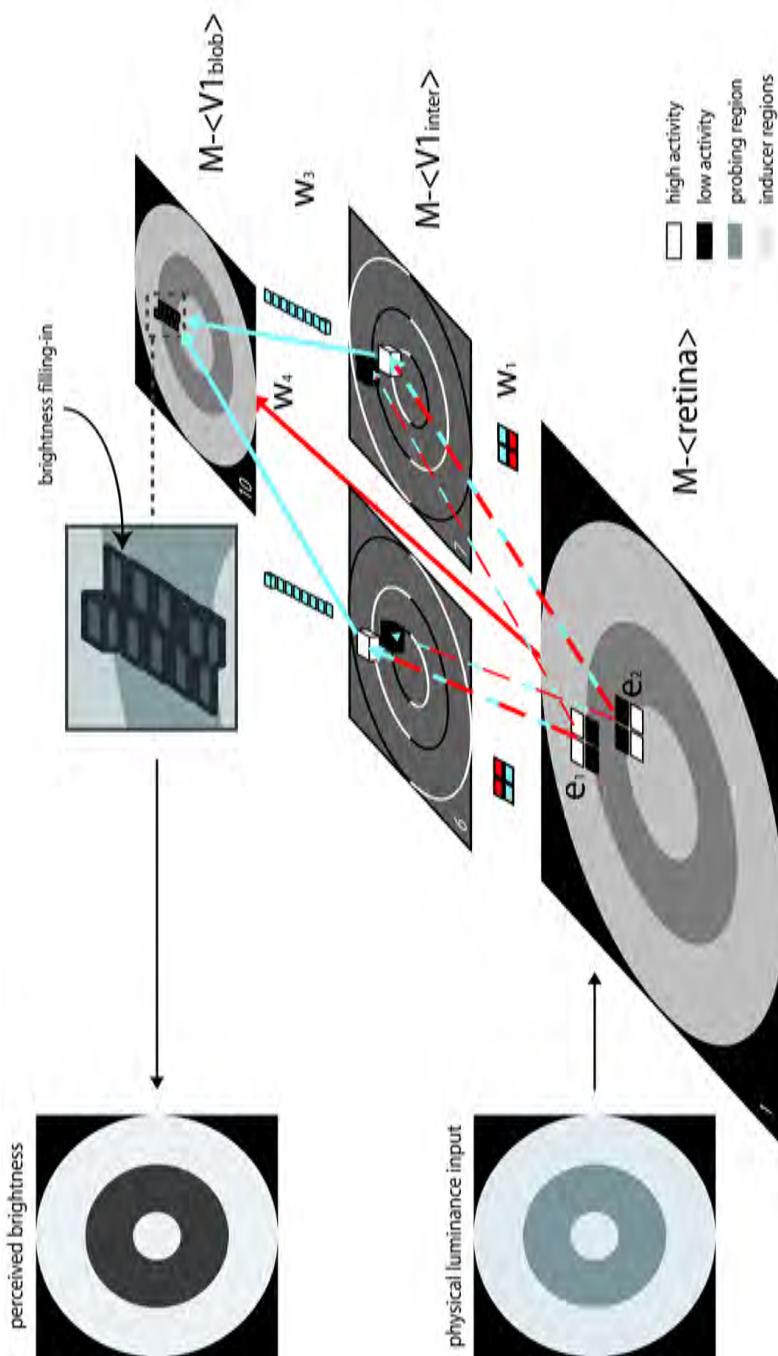


## 4 Brightness Induction Modeling

**Figure 2.** The neural model and MRI data conjointly represented in a common brain space framework. In this framework, neuromodeling and neuroimaging data can be directly integrated. On the left the CBS is shown similar to its visualization in Neurolator, whereas the right shows a schematic outline of how we connected the CCUs to this irregularly shaped patch of vertices. Network-Brain links (NBLs; 1) connect network CCUs with corresponding positions on the cortical sheet. In this example, the NBLs connect the upper quadrant of model area M-<V1<sub>blob</sub>> (2) to regions in the left visual cortex (3) representing the inducer (orange and blue NBLs/cortical patches) and probing (green NBLs/cortical patches) areas as identified by an fMRI localizer mapping. To optimally visualize the relation between V1 and M-<V1<sub>blob</sub>>, the cortical sheet is inflated (inset) and flattened. However NBLs can be connected to all sorts of cortical mesh representations. Note that this example shows connections from initial tests, modeling input to a stimulus with different annulus widths than used in the final simulations.

cells at a surface boundary will mainly excite or inhibit the corresponding M-<V1<sub>inter</sub>> cell if the excitatory connections represent the lighter (as in e<sub>1</sub>) or darker (e<sub>2</sub>) surface, respectively. In this figure, the M-<V1<sub>inter</sub>> layers with top (left sheet) and bottom (right sheet) contour detection cells are shown, but activation mechanisms are similar for the other six layers. The top and bottom contour detection layers have opposite activity profiles, which will lead to antipodal activations at the subsequent stage.

In the second stage, this contour-related activity modulates luminance-related activity in M-<V1<sub>blob</sub>>. CCUs of M-<V1<sub>inter</sub>> with high activity (the white cubes) will inhibit corresponding cells in the next layer according to the depicted connection pattern. Thus, active CCUs in the top contour detection (left) sheet will induce a downward inhibition, whereas active CCUs in the bottom contour detection (right) sheet induce an upward inhibition. This results in an inhibition of the probing region, corresponding to the decreased perceived brightness of this region. However, although a gray area appears to be darker in bright surroundings, the illusion also holds true when abutting inducers become darker, causing the grey probing region to appear brighter. Our network is also able to handle this stimulus situation, as half of the M-<V1<sub>inter</sub>> layers excite rather than inhibit M-<V1<sub>blob</sub>> activity. As shown in Figure 1, layers 2-5 of M-<V1<sub>inter</sub>> are identical to layers 6-9, except for the polarity of their connection to M-<V1<sub>blob</sub>>. Therefore, when inducers become darker, active CCUs in layers 2 and 3 will



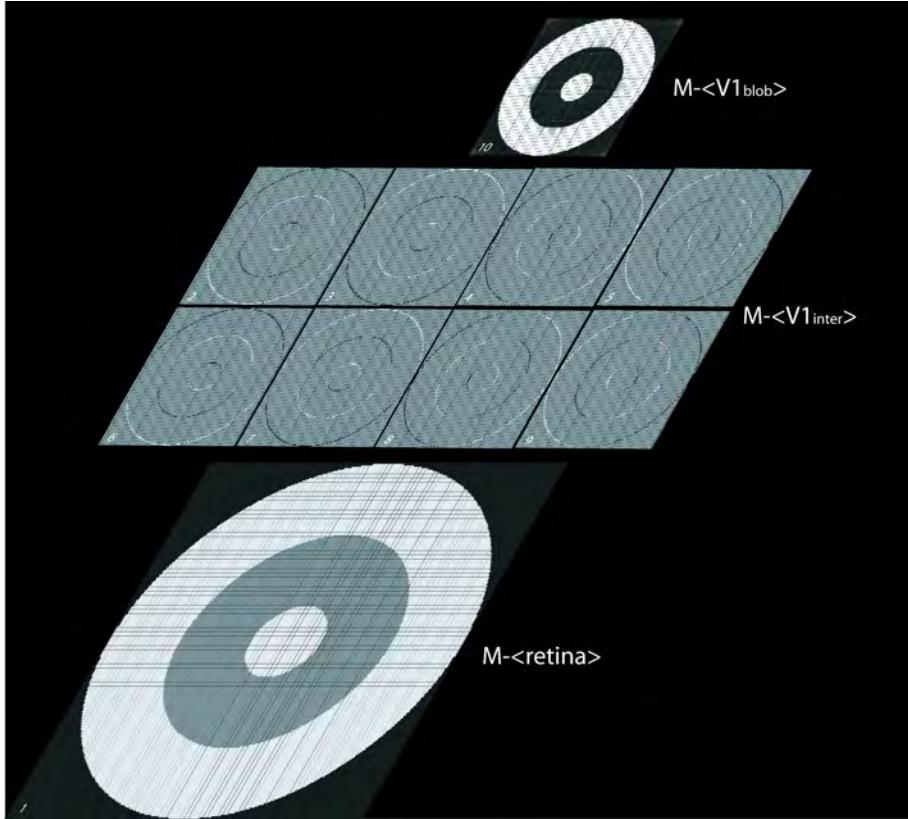
## 4 Brightness Induction Modeling

**Figure 3.** Illustration of the model's computational mechanisms. The model areas display schematic activation profiles (grey levels indicate activity strength) with several CCUs shown magnified on top to illustrate the activity flow. Physical luminance information from the stimulus (with a probing region surrounded by bright inducers) enters M-<retina>. This information is further processed in two discrete stages: In the first stage, contour information is extracted from local luminance discontinuities, which is signaled from M-<retina> to M-<V1<sub>inter</sub>>. Activity in M-<V1<sub>inter</sub>> depends on the contour detection sensitivity, specified by the connection pattern (w1). An M-<V1<sub>inter</sub>> CCU that is connected to four M-<retina> CCUs representing a homogeneous surface, will receive a balanced input of excitation and inhibition. In contrast, cells at a surface boundary will mainly excite or inhibit the corresponding M-<V1<sub>inter</sub>> cell if the excitatory connections represent the lighter (e1) or darker (e2) surface, respectively. In this figure, the M-<V1<sub>inter</sub>> cell clusters with top (left sheet) and bottom (right sheet) contour detection are shown, but activation mechanisms are similar for the other six contour detectors. Note that the top and bottom contour detection layers have opposite activity profiles. In the second stage, activity reflecting physical luminance (from M-<retina> via w4) is integrated with contour-related activity (from M-<V1<sub>inter</sub>> via w3) in M-<V1<sub>blob</sub>>. M-<V1<sub>inter</sub>> CCUs with high activity (the white cubes) will inhibit corresponding cells in the next layer according to the depicted connection pattern (w3). Thus, active CCUs in the top contour detection (left) sheet will induce a downward inhibition, whereas active CCUs in the bottom contour detection (right) sheet induce an upward spread. This resulted in an inhibition of the probing region, corresponding to the 20 decreased perceived brightness of this region (see inset). The thickness of the arrows indicates the strength of activity conveyed. Further conventions are identical to Figure 1.

increase activity in the probing region of M-<V1<sub>blob</sub>>, as opposed to the inhibitory influence of layer 6 and 7 as shown in the example in Figure 3.

### **Activity in M-<retina> and M-<V1<sub>blob</sub>> reflect physical and perceived brightness respectively**

Figure 4 shows an example of the model's behavior. As in Figure 3, activity in M-<retina> reflects the physical luminance of the image, which provides information on luminance discontinuities to model area M-<V1<sub>inter</sub>>. The activity profile of layers two to nine, constituting M-<V1<sub>inter</sub>>, emphasizes



**Figure 4.** Activity in the network. When luminance values of a stimulus with bright inducers (see lower left image in Figure 3) enter the network, activity in model area M-<retina>, M-<V1<sub>inter</sub>>, and M-<V1<sub>blob</sub>> reflect physical brightness, contours, and perceived brightness, respectively. Grey values indicate response strength (black = no activity, white = maximum activity). See Figure 1 and 3 for network details.

the contours of the surfaces, whereby the distribution of low (black) and high (white) activations depend on the contour detection sensitivity of the specific layer. This contour information subsequently modulates activity in layer M-<V1<sub>blob</sub>>, leading to decreased activity in the probing region.

Figure 5 provides the opportunity to directly compare the physical and perceived luminance, as reflected by M-<retina> and M-<V1<sub>blob</sub>> respectively. Here, we show the network's behavior for two stimuli at the extremes of the continuum (i.e., at the maxima of the sinusoidal changes), whereas Figure 6 gives more information on the model's performance along the entire continuum. When comparing activity in the probing region of M-<retina> and M-<V1<sub>blob</sub>> in the first row of Figure 5, one can clearly see that activa-

## 4 Brightness Induction Modeling

tion is reduced in the second area, corresponding to a decreased brightness percept. The opposite is true for the second row, in agreement with an increase in perceived brightness of a gray surface in dark surroundings. So, although the equal luminance of the probing region in both stimuli elicited an identical activation in M-<retina>, the activation in M-< $V1_{blob}$ > greatly differs, due to contextual influences.

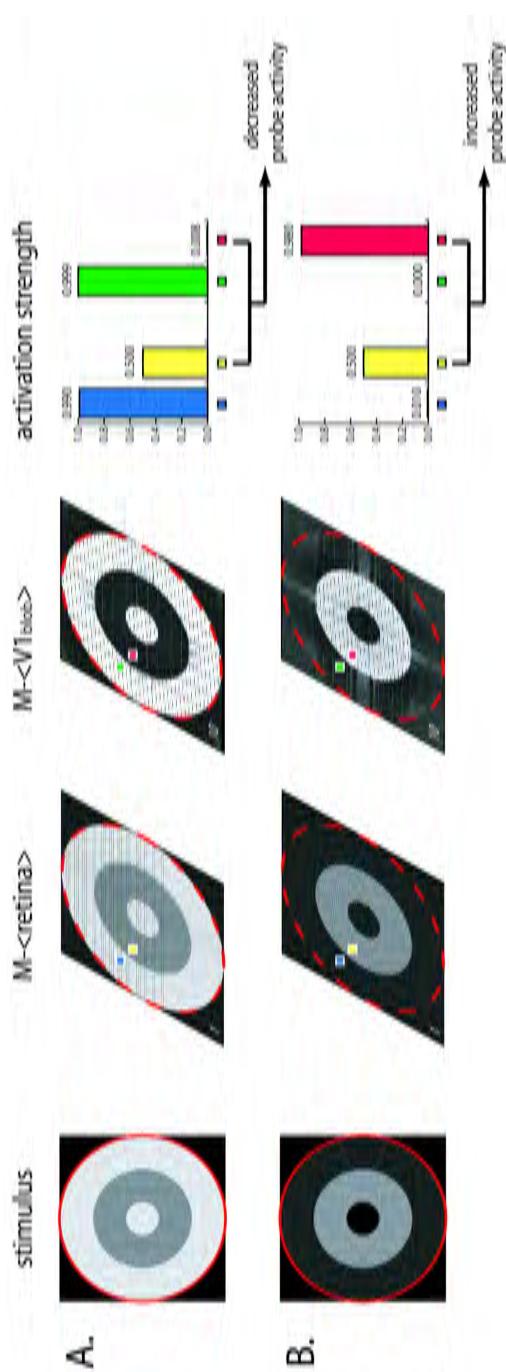
### **Activity in M-< $V1_{blob}$ > corresponds to hemodynamic correlates of perceived brightness in V1**

The NBLs in the CBS framework provide a precise one-to-one correspondence between the CCUs and the cortical patches representing the inducer and probing regions. Moreover, they allow the direct projection of network activity on the cortical sheet.

We connected the cortical mesh to a running neural network, and created time series of M-< $V1_{blob}$ > activity changes that were projected on the corresponding cortical representation (see Figure 2) Figure 6 displays the average time series of the simulated data in the localized probe and inducer POIs. As shown in Figure 6a and 6b, the synaptic activity (netiabs) in the inducer area is clearly in-phase with the actual luminance changes of the inducer. In contrast, activity changes in the probing region are in anti-phase, robustly mimicking the perceived brightness changes. The second column of Figure 6 presents the simulated fMRI signal (ynoise), which is obtained by convolving the spike data with an HRF function and by adding Gaussian noise.

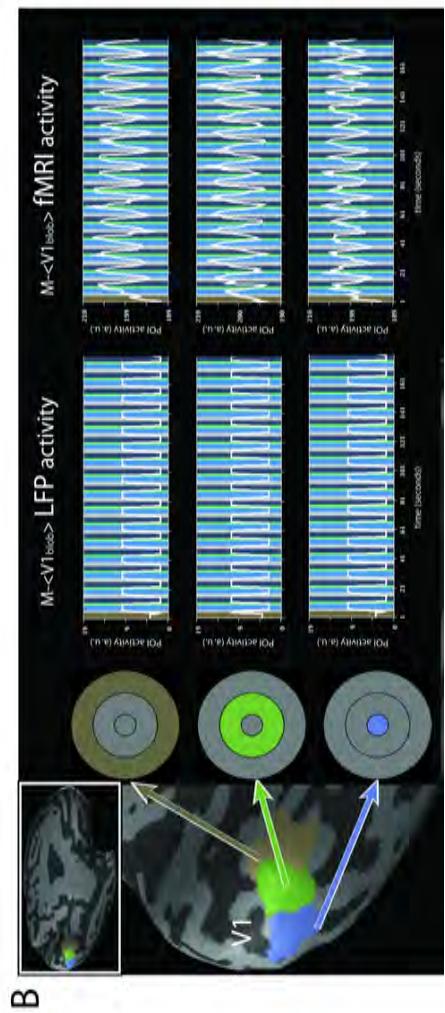
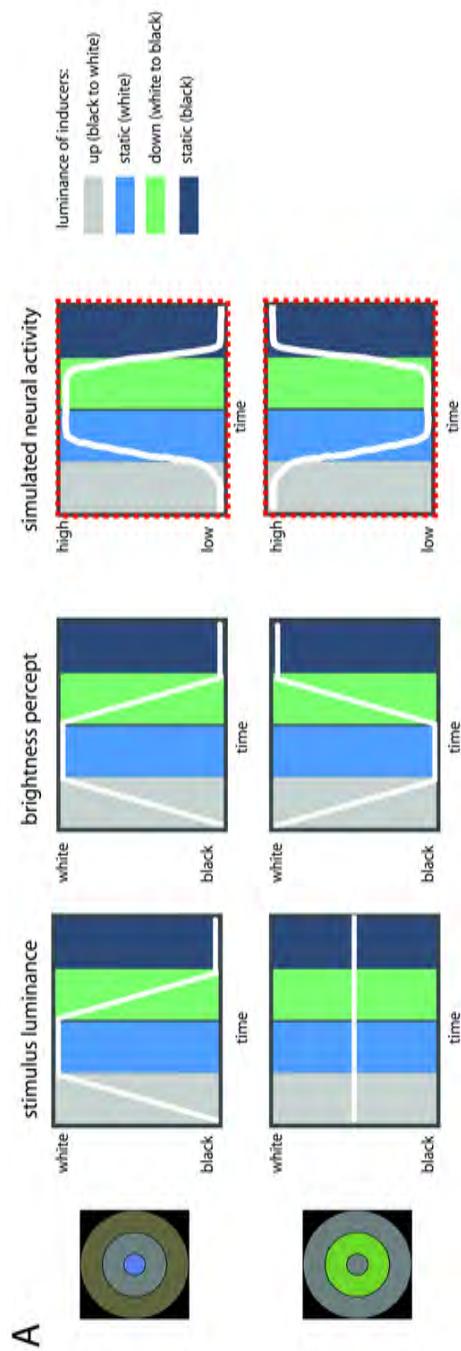
Figure 7 shows the event-related averages of this signal in the illusory (left column) and control (right column) condition when the luminance of the inducers gradually increase ("up"; green line) or decrease ("down"; blue line). In line with the activity patterns in Figure 6b, the event-related fMRI signals of the up and down period are in anti-phase in the inducer and probe POI. This anti-correlation between inducer and probe POI activity also emerges in the statistical map: As shown in Figure 8, voxel-wise regression clearly segregates the inducer and probe POIs (compare Figure 8A and 8B), indicating that when activity increases in the inducer POI, it decreases in the probe POI and vice versa.

This anti-phasic activity of inducer regions following luminance changes, and of probe regions corresponding to perceived luminance, is very obvious when the amplitudes of the event-related averages are plotted, as is done in Figure 9: In the "up" condition (green bars), inducer POI activity goes up (in line with increasing inducer luminance), whereas probe POI activity goes



#### *4 Brightness Induction Modeling*

**Figure 5.** Activity profiles in M-<retina> and M-<V1<sub>blob</sub>> induced by a stimulus with bright (A.) versus dark (B.) inducers. A.) Activation in the probing region is reduced in M-<V1<sub>blob</sub>> compared to M-<retina>, corresponding to a decreased brightness percept.B.) Probe activity is higher in M-<V1<sub>blob</sub>> than M-<retina>, in agreement with an increased perceived brightness of the probing region. Comparison of the red and yellow bars, indicating activation strength of the corresponding CCUs, confirms that CCUs representing the probing region elicited an identical activation in M-<retina>, whereas these activations diverge in M-<V1<sub>blob</sub>>. Note that the (dashed) red circle is for illustration purposes only.



#### 4 Brightness Induction Modeling

**Figure 6.** Simulated neural activity elicited by dynamically changing inducer luminance. A. Illustration of experimental design. Each trial starts with an "up" event in which inducer luminance changes from black to white, accompanied by an increasing perceived brightness of the inducers, and a decreased brightness percept of the probe annulus. After a static period, there is a "down" event in which inducer luminance decreases with positive and negatively correlating brightness percepts of inducers and probe regions, respectively. Our simulations (right column) suggest that neural activity in blob areas of V1 change in correspondence to the brightness percept (middle column). Thus, activity in the inducer (upper row) and probe (lower row) regions show anti-phasic changes in activity. B. The average time series of probe and inducer POIs, respectively reflecting activity of probe and inducer CCUs in model area M- $\langle V1_{blob} \rangle$ . The brown, green, and blue patches (POIs) on the inflated mesh (medial view on the occipital pole) correspond to the cortical representations of the outer inducer, probe, and inner inducer in the primary visual cortex, respectively. Note that in this and in the subsequent figures, the colored surfaces in the accompanying stimuli indicate the corresponding POIs rather than the actual stimulus colors. The left column shows the spike output of the model in these three POIs, whereas the right column presents the simulated fMRI signal. The timeseries show the antiphasic activity changes in inducer and probe POIs (compare for example, activity in the first and second timeseries during the blue interval).

down (though probe luminance is fixed). The reversed is true in the "down" condition (blue bars), showing that increased as well as decreased illusory brightness changes are reflected by hemodynamic activity. This reversal effect of probing region activity was absent in the control condition (i.e., fixed black probing region). Moreover, Figure 9 provides a direct comparison between POI activity of simulated and empirically-obtained fMRI data. In the first case, POI activity represents data of probe or inducer CCUs in model area M- $\langle V1_{blob} \rangle$ , which are projected to the corresponding mesh vertices by NBLs (see Figure 2 and Methods for details). In the second case, POI activity reflects activity of real fMRI measurements which were aligned to the corresponding anatomical cortical vertices. The similarity of the amplitude patterns in the left (simulations) and right (empirical data) plots of Figure 9 suggest that M- $\langle V1_{blob} \rangle$  activity can predict observed activity in human V1 fairly well.

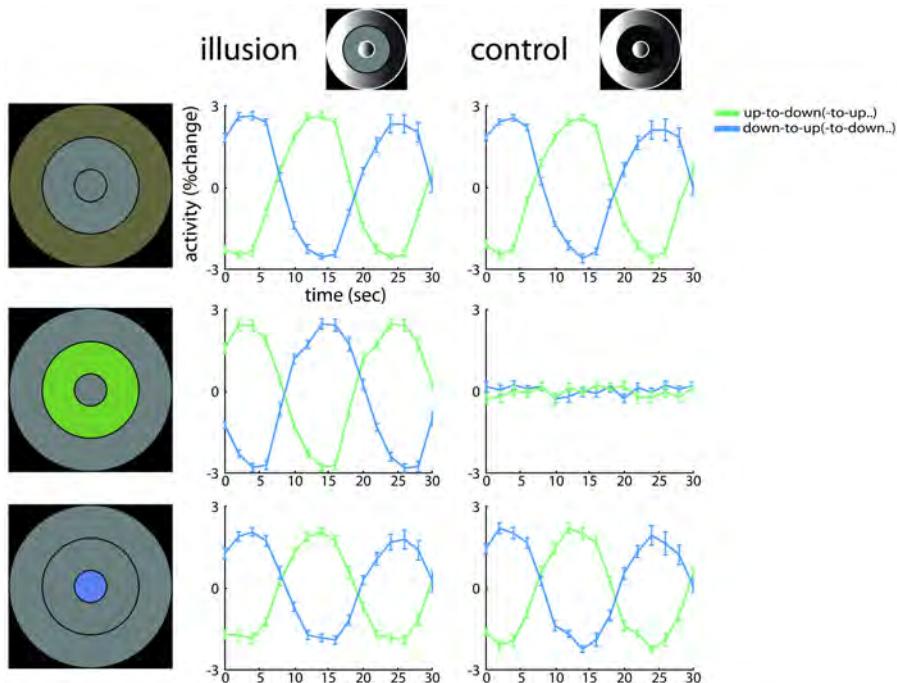
## **Discussion**

In this paper, we apply a new approach to compare empirical data and model predictions in a common brain space (Goebel & Horwitz, in preparation; Goebel & De Weerd, 2009). We suggest that this approach can facilitate hypothesis testing, thereby contributing to deciphering cortical circuits involved in surface brightness perception and other perceptual and cognitive abilities. Below, we discuss the potential neural mechanisms of brightness perception and how we and others have modeled these mechanisms. Finally we discuss current possibilities of the CBS framework as well as future perspectives.

### **Neural mechanisms of brightness perception**

At a subcortical level, the neural mechanisms of contour and surface perception are relatively well understood: The visual image is filtered by center-surround antagonistic RFs of retinal ganglion and Lateral Geniculate Nuclear (LGN) cells, leading to image normalization. Since normalization leads to a strong emphasis on discontinuities in luminance distribution, the primary visual cortex (V1) mainly receives contour- rather than surface-related information. Most models suggest that at the cortical level surfaces have to be reconstructed from local contour information (Grossberg, 1987a,b; 2003), which can also explain why brightness percepts can differ from the actual brightness in brightness illusions. However, not all surface information is lost in the subcortical processing stage: The LGN (Rossi et al., 1996), V1 blobs (Roe et al., 2005), and V2 thin stripes (Ts'o et al., 2001; Wang et al., 2007) contain neurons that are readily stimulated when large, homogenous stimuli are placed over their RFs. The presence of such 'Luxotonic' (Kayama et al., 1979) or 'Type-I' (Kinoshita and Komatsu, 2001) neurons that code absolute luminance, was confirmed by others in monkey V1 ( Friedman et al. 2003; Huang and Paradiso, 2008; Kayama et al., 1979; Kinoshita and Komatsu, 2001; Maguire and Baizer, 1982; Wachtler et al., 2003) and in cat area 17 (Bonhoeffer et al., 1995; Shoham et al., 1997). Visual illusions offer a valuable tool to study at which stage of processing surface percepts are constructed, as physical and perceived surface features (e.g., brightness or color) can be segregated. The findings above suggest that physical surface information is reflected along the subcortical and cortical pathway, but it is less clear at which stages of processing perceived surfaces are represented. LGN activity appears to only reflect physical brightness (Rossi et al., 1996), whereas divergent observations are reported for V1

#### 4 Brightness Induction Modeling

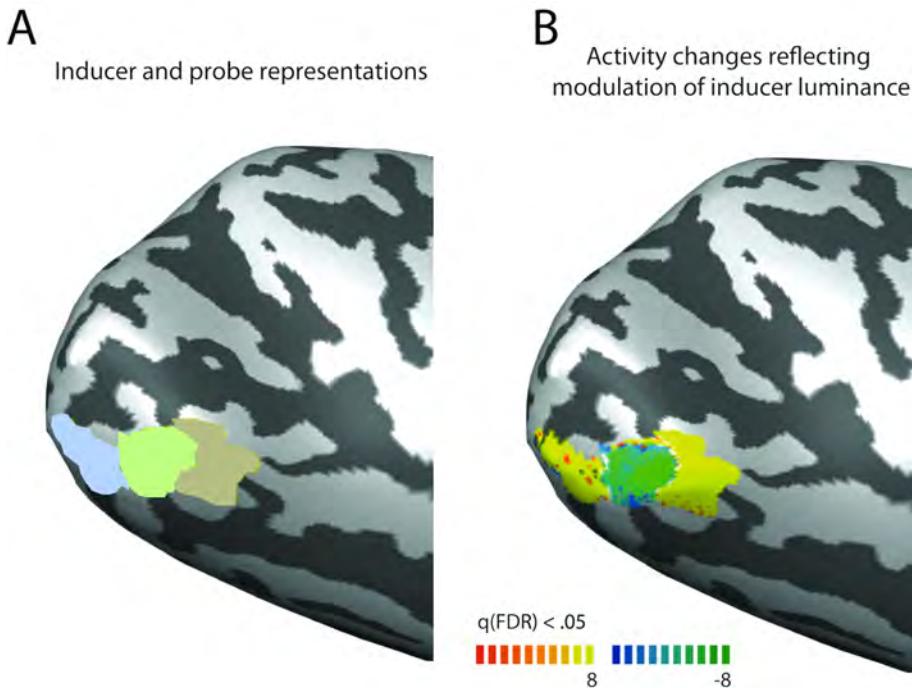


**Figure 7.** Figure 7. Event-related averaged time series of the simulated fMRI data for the outer inducer (upper row), probe (middle row), and inner inducer (lower row). The left and right columns show the event-related averages of the illusory and control condition, respectively. The probe had a static intermediate luminance in the illusory condition, whereas it was black in the control condition. In both conditions, the inducers gradually varied in luminance from black to white and vice versa (see Figure 6b), as indicated by the gradient in the column header icons. The profile of the event-related averages, as well as the differences between conditions, is akin to the event-related averages of the observed fMRI signal (not shown). Note the sign-reversal of the probe POI average, in comparison to the inducer POIs in the illusion condition, whereas there is no probe POI activity in the control condition. Errorbars indicate 1 standard error of the mean (SEM).

which might be due to the different visual illusions that were used. Studies investigating the Craik-O'Brien-Cornsweet illusion (Cornsweet, 1970) observed mainly perceived brightness correlates in V2 (Hung, Ramsden, Chen, and Roe, 2001; Perna, Tosetti, Montanaro, and Morrone, 2005; Roe et al., 2005), although V1-V2 interactions significantly contribute (Hung, Ramsden, and Roe, 2007; and see Boyaci, Fang, Murray, and Kersten, 2007 for cornsweet effects in V1). In contrast, activity related to neon color spreading has been established in monkey V1 (Murakami et al., 1997; Matsumoto & Komatsu, 2005; Komatsu, Kinoshita, & Murakami, 2000) and human V1 (Sasaki and Watanabe, 2004). Also for brightness induction, which is the type of brightness illusion used in the present study, correlates of perceived brightness have been observed in V1 (Pereverzeva and Murray, 2008; van de Ven et al., submitted; but see Cornelissen et al., 2006) in V1. Together, these findings suggest that, at least for brightness induction, both physical as well as perceived brightness-related surface information is represented in V1. Note that neural activity underlying brightness perception can be modulated by cognitive factors. For example, Sasaki and Watanabe (2004) showed that neural correlates of color spreading are constrained to V1 when attention is controlled, whereas otherwise correlates of spreading effects can be found in extrastriate areas as well. Surprisingly, Mendola et al (2006) observed decreased activity in V1 and V2 during luminance filling in. This discrepant finding might be related to the different paradigm used in that study (i.e., Troxler's fading; Troxler, 1804), which might have engaged higher-level processes (indicated by the additional activations observed in cortical association regions).

Contour and surface information appear to be processed in two segregated, but interacting, streams: Neurons in V1 blobs respond to homogenous surfaces, whereas V1 interblob regions contain neurons that are preferentially stimulated by oriented contrasts (e.g. contours and edges). Likewise, neurons in V2 thin stripes areas react to surfaces (Roe et., 2005), while V2 interstripe 13 regions show preferential responses to lines and edges (Bartfeld and Grinvald, 1992; Blasdel et al., 1985; DeYoe and Van Essen, 1988; Fitzpatrick et al., 1985; Hubel and Livingstone, 1987; Landisman and Ts'o, 2002; Lu and Roe, 2008; Yabuta and Callaway, 1998). Taken together, these data support the idea that there are separate (but interacting) processing streams for contour and surface processing, with a blob/thin stripe system for surface representation and an interblob/interstripe system for contour representation. For a more extensive review on neural mechanisms of brightness perception, the reader is referred to the existing literature

#### 4 Brightness Induction Modeling



**Figure 8.** A. Medial view on the occipital pole of an inflated cortical mesh. The brown, green, and blue patches (POIs) correspond to the cortical representations of the outer inducer, probe, and inner inducer in V1, respectively. B. Results of contrasting V1 activity (with the map constraint to the V1 POIs) of the network simulations that is in-phase (red-yellow) and in counter-phase (blue-green) with the changing luminance of the inducers. Note that the borders where in-phase and counter-phase activity segregate (in B), coincide with the borders between the probe and inducer regions (in A).

(Pessoa and De Weerd, 2003; Goebel and De Weerd, 2009; Neumann et al., 2001).

#### Our model and other computational models of brightness perception

Our model is based on empirical research discussed in the previous section. An essential feature in our model's architecture is the modulation of physical brightness information represented by "luxotonic" cells in  $M-, by contour information in  $M-. Results showed that our model performs well in simulating dynamical changes in perceived surface brightness. It was able to predict illusory brightness changes in a grey area of constant luminance when brightness of its surroundings was dynamically$$

varied. Moreover, the activation changes in CCUs representing the probing region, varied in a similar fashion as V1 activity in a subject presented with the same stimulus design. This suggests that V1 plays a role in generating surface brightness percepts, which is not only in agreement with our own fMRI data (van de Ven et al., submitted) and those of others (Pereverzeva and Murray, 2008), but also with neurophysiological findings in the cat (Rossi et al., 1996; 1999).

Our modeling approach is rather different from previous modeling approaches. Our work started from the assumption that surface brightness is explicitly represented in early visual cortex, represented by the “luxotonic” cells in model area V1 signaling surface information. This hypothesis differs from proposals that surface perception relies on ‘symbolic tagging’ (Dennett, 1991). According to the symbolic tagging theory, surface qualities are not encoded in retinotopic areas. Instead, surfaces are implicitly (symbolically) represented as a lack of discontinuities in between identified surface boundaries, and surface encoding takes place entirely at later visual stages, where surfaces form an integral aspect of object representations. In addition, our surface representation at the earliest cortical stages of visual processing is also not in agreement with multi-scale spatial filtering theories assuming that brightness perception does not require an explicit spreading mechanism (Blakeslee et al., 2005; Dakin and Bex, 2003; McCourt, 1982; Purves et al., 1999; Stromeier et al., 1984). Based on multi-scale spatial filtering theory, some models were developed to simulate brightness assimilation (Barkan, Spitzer, and Einav, 2008; Blakeslee and McCourt, 1999; 2001; 2004; Otazu, Vanrell, and Párraga, 2008), such as in the classic White’s effect (White, 1979). The most well-established model in this category is the oriented difference-of-Gaussians (ODOG) model of Blakeslee and McCourt (1999, 2001, 2004) that performs an oriented multiscale spatial filtering of input and a subsequent global contrast normalization to equalize responses at each orientation across the visual field. This model can simulate a range of brightness illusions (Blakeslee and McCourt, 2004). Although we do not model receptive fields (RFs) with oriented multiscale filters, the RFs in our model are orientation-specific. In addition, these RFs have excitatory and inhibitory subzones, which effectively resemble local normalization. Yet, although RF creation in the current model thus already in some sense share similarities with the approach of Blakeslee and McCourt (2004), we will implement Gabor filters (Jones and Palmer, 1987) in future versions of our model to simulate V1 responses even more accurately. Although differences of Gaussians mimic retinal ganglion responses properly (Rodieck, 1965)

## *4 Brightness Induction Modeling*

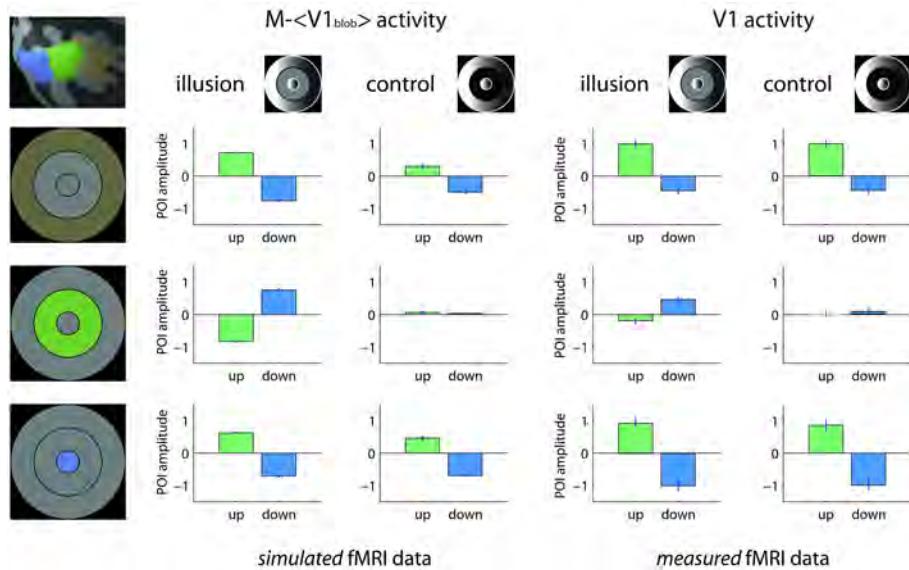
other aspects of the Blakeslee and McCourt (and related) models have a less close neurobiological correspondence, and some processes such as the global normalization across orientations might be neurally less plausible as normalization is commonly considered a more local process. Interestingly, a recent neuroimaging study of Kay et al. (2008) used multiscale spatial filtering to decode natural scene processing in visual cortex measured. In the first stage of the fMRI experiment, they recorded fMRI responses in V1-V3 to 1,750 natural images and used these data to estimate a RF model (based on a Gabor wavelet pyramid) for each brain voxel. Subsequently, this model was employed to predict neural responses to 120 novel images and then to identify which image was viewed by finding the best match between model predictions and observed fMRI activity (Fig. 1b). We believe that our CBS framework provides an excellent context to integrate the modeling efforts like those of Blakeslee and McCourt with advanced neuroimaging analyses such as those employed by Kay et al. (2008).

Our model has some commonalities with the models of Grossberg (Grossberg, 1987a, 1987b; 2003; Grossberg and Todorovic, 1988; Grossberg and Raizada, 2000; see Neumann et al., 2001 for a related approach) in which surfaces are created by a spreading activation of surface features from the contours of the surface inward. However, according to Grossberg's theory, contours initiate surface spreading and all surface information is reconstructed by an interaction between contour information and surface features. Different to our approach, these surface features are sampled from positions with large luminance or color transitions, to discount variable illumination. In our model on the other hand, we explicitly model luminance-related information of the entire surface, based on direct input (modeled as w4) from the subcortical pathway (i.e., cells in the LGN signaling surface brightness; Rossi and Paradiso, 1999). Subsequently, the surface information represented by these "luxotonic" CCUs is modulated by – rather than reconstructed from – contour information. Arrington (1994) used a simplified version of Grossberg's model to simulate psychophysical findings in the masking experiments of Paradiso and Nakayama (1991), which tested the speed of brightness spreading in a homogeneous bright circle on a black background, by disrupting the inward spread with a masking stimulus. Arrington's model contained only four neuronal layers: an On-FCS and an Off-FCS layer modeling retinal cells with DOG RFs, and one BCS layer in which "boundary cells" are activated at locations where On-FCS and Off-FCS activations are spatially adjacent. Finally, there is the filling-in layer where brightness spreading by FCS input is regulated by parallel BCS input. This simple but elegant

model was able to simulate, with a range of masks, the brightness percepts observed by Paradiso and Nakayama (1991). The model of Arrington is, to our knowledge, the only other computational model that models brightness induction using a homogenous stimulus (which can be compared to our homogenous probe region). However, whereas Arrington's model aims to simulate the rapid (i.e., in millisecond-range) brightness filling-in in a static stimulus with fixed luminance, our model attempts to simulate dynamic changes in fMRI activity induced by slow (i.e., in second-range), continuous luminance changes in a dynamic stimulus. These different objectives led to different modeling choices, making it difficult to directly compare our model to that of Arrington or, in fact, to all other models as none of them used –to our knowledge– dynamically changing stimuli.

As the present paper represents a 'proof of principle', we deliberately chose a simple network architecture that focuses on a limited set of mechanisms at a large-scale level. Although it is interesting to see that this limited model has sufficient features to mimic empirically observed fMRI activity, it is clear that we did not implement many relevant principles of neural processing in early visual areas, such as orientation and spatial frequency tuning, or processing in pre-striate processing pathways. Furthermore, although our and other models (e.g., Heitger et al., 1998) show that appropriate contour detection can be obtained with solely feedforward and horizontal connections, boundary detection can be performed faster and more reliable by adding feedback connections (Jehee, Lamme, and Roelfsema, 2007; see also Roelfsema et al., 2002). Finally, we modelled only V1 as we (van de Ven et al., submitted) and others (Pereverzeva and Murray, 2008) observed the strongest fMRI correlates of brightness induction in this area. However, V2 appears to play a significant role in computing perceived brightness in the Cornsweet stimulus (Roe et al., 2005). This suggest that both V1 and V2, and especially their interactions (Hung, Ramsden, and Roe, 2007), are important for generating brightness percepts, although their respective contributions seem to shift depending on the studied brightness illusion. Computational modeling offers a promising way to gain insights in the roles of V1 and V2 in brightness perception. By extending our model with recurrent V1-V2 loops, and subsequently comparing observed and simulated fMRI data at V1 and V2 level in different brightness illusion paradigms, a general mechanism of brightness computation might be uncovered that reconciles the various findings.

#### 4 Brightness Induction Modeling



**Figure 9.** Comparison between amplitudes of the simulated (left) and real (right) fMRI signals in probe and inducer POIs. Amplitudes were quantified by subtracting an early time segment (2-4 s) from a late time segment (8-10 s) of the event-related averages, and averaging the differential responses. Amplitudes are shown when inducer luminance increases (UP) and decreases (DOWN) in the outer inducer (upper row), probing region (middle row) and inner inducer (bottom row). For the illusion condition, both the simulated and real fMRI signal show inducer activity in phase with the physical luminance changes, but probing region activity in anti-phase with luminance changes. This reversal effect is not found in either the simulated or real signal in the control condition. Amplitudes are normalized to the maximum amplitude across POIs. Errorbars indicate 1 SEM.

## **Large-scale neural models and neuroimaging data in a common brain space: Current possibilities and future perspectives**

The approach used in the present paper shows that empirical data and model predictions can be compared in the same brain space. Importantly, when simulated and measured data co-exist in the same representational space, the same analysis tools (e.g., univariate GLM, multi-voxel pattern analysis, effective connectivity analysis) can be applied to both data sets. Furthermore, dynamic activity patterns of running neural network models can evolve directly on cortex representations or meshes obtained from MRI scans of specific subjects (Kriegeskorte and Goebel, 2001; Sereno et al., 1995), and therefore, online hypothesis testing against incoming empirical data is possible.

In addition to the manner in which neuromodeling and neuroimaging data were compared in the CBS framework in the present paper, there are several more useful approaches that have not been shown in the present paper. We have confined our comparison to modeled activity in two V1 model-layers, which was projected as predicted activity in a single (functionally localized) anatomical region in the human brain (V1). However, comparisons between multiple layers and multiple regions are also possible. In addition, comparisons do not have to be constrained to the visual cortex, but can cover different portions of the brain. To study for example the observations of Sasaki and Watanabe, 2004, the model could be extended with layers representing the frontal eye fields (FEF) and regions in the posterior parietal cortex exerting top-down attentional influences (Corbetta and Shulman, 2002) on brightness computations in the visual areas. Subsequently, we could study how the processing of brightness illusions is modulated by attentional demands in an fMRI and a simulation study. In our CBS framework, we would then compare concurrent activity in higher-level attentional control regions (e.g., functionally localized FEF; Tehovnik et al., 2000) and visual areas (e.g., V1) that is observed in the empirical and in the modeling study.

In general, NBLs can connect output of different types of neural nets to any kind of functionally localizable regions, ranging from sensory retinotopic (DeYoe et al., 1996), tonotopic (Formisano et al., 2003), and somatotopic (Blankenburg et al., 2003) areas, to higher level regions such as object-preferring higher visual areas (e.g., fusiform face area; Kanwisher et al., 1997) or language areas (Fernández et al., 2001). Likewise, all sorts of anatomically defined regions can be linked to model areas. If a localizer

#### *4 Brightness Induction Modeling*

scan had been lacking, we could have for example projected our V1 activity to an anatomically defined V1 (Hinds et al., 2009).

Further, not only perceptual processes, but also potential neural substrates of cognitive processes (e.g., Brown and Braver, 2005; O'Reilly, 2006; Tagamets and Horwitz, 1998) can be facilitated by the CBS framework. The wide applicability of integrating simulated and real data in 17 brain space is becoming increasingly recognized. Arbib and coworkers (1995) were the first to compare computational simulations (on saccade generation) to neuroimaging (Positron Emission Tomography or PET) data. Subsequently, Tagamets and Horwitz (1998) simulated fMRI data on a visual short-term memory task by transforming spiking output from their large-scale neural model into fMRI activity by means of a hemodynamic response function. Furthermore, they downsampled the resulting time series according to the appropriate TR and mimicked features of the empirical study, such as trial repetition and corresponding intertrial intervals. The Tagamets and Horwitz model was extended to working memory processes in the auditory domain by Husain and colleagues (2004), and for the first time a quantitative comparison between simulated and observed fMRI data was made. Other modeling efforts that simulate fMRI data in the field of working memory come from Deco and coworkers (Deco et al., 2004), and from Chadderan and Sporns (2007) who implemented dopaminergic modulations in their neural network. Although working memory received a lot of attention, there have also been attempts to relate neuromodeling and fMRI data in the field of attention (Corchs and Deco, 2002) and conflict processing (Herd, Banich, and O'Reilly, 2006; Brown and Braver, 2005).

However, none of the previously mentioned integration approaches provided the direct one-on-one topographic correspondence between computational units at the model and brain level that our CBS framework offers. This close correspondence greatly facilitates the direct spatial comparison between the simulated and the fMRI results, and it permits modeling the contribution of anatomical microcircuits to predicted fMRI activity. This permits investigators to develop increasingly more mechanistic and biologically valid types of modeling, which may stimulate a tighter integration of computational and experimental neuroscience. Specifically, our topographically more detailed modeling approach could assist in bridging the theoretical and empirical gaps between multiple levels of resolution of the various measurement modalities. That goal is advanced further by the ability of our modeling approach to generate multiple output data types, ranging from integrative synaptic activity levels and mean spike activity to

fMRI and EEG/MEG data types. In CBS, it is possible to simply switch between displaying the different output data types produced by the same running network simulation. This switching between data types in a common space can be easily accomplished as the CBS modeling framework can be conveniently integrated with recently developed extensions of the Brain-Voyager QX software package that permits signal projection for fMRI as well as electro- & magneto-encephalography (EEG/MEG) data (Esposito et al., 2009).

In addition to integrating levels of understanding at different spatial (and temporal) resolution, the detailed modeling approach we propose might help to reconcile apparently conflicting results from human fMRI studies and spike recording studies in animals that investigate the same phenomena across species (Maier et al., 2008). For example, we believe that the ability to compare different model output in terms of different methods may help to investigate unanswered fundamental question about the coupling between integrated synaptic activity and fMRI signal (Logothetis et al., 2001), or about the relationship between neuronal activity distributions and the spread of fMRI signal.

With the advent of high field fMRI imaging ( $\geq 7$  Tesla), it may become possible to image brain activity in response to perceptual and cognitive events at a high spatial resolution that may enable investigators to monitor activity in separate layers, columns, and other anatomical and functional subcompartments of cortical areas in the human brain. The activity patterns predicted by advanced computational models that implement the details of cortical circuitry may therefore be projected onto a very detailed brain space, in which also the empirical data can be collected. For example, it may become possible to directly test the hypothesized specific contributions of different compartments of V1 and V2 to surface brightness perception. We suggest that the approach of testing model predictions in brain space will become a particularly important tool in future cognitive neuroscience research.

## References

- Arbib, M.A., Bischoff, A., Fagg, A.H., Grafton, S.T., 1995. Synthetic PET: Analyzing Large-Scale Properties of Neural Networks. *Human Brain Mapping*, 225-233.
- Arrington, K. F., 1994. The temporal dynamics of brightness filling-in. *Vis Res* 34, 3371-3387.
- Bartfeld, E., Grinvald, A., 1992. Relationships between orientation-preference pinwheels, cytochrome oxidase blobs, and ocular-dominance

#### *4 Brightness Induction Modeling*

- columns in primate striate cortex. *Proc Natl Acad Sci U S A* 89, 11905-11909.
- Blakeslee, B., McCourt, M.E., 1999. A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Res* 39, 4361-4377.
- Blakeslee, B., McCourt, M.E., 2001. A multiscale spatial filtering account of the Wertheimer-Benary effect and the corrugated Mondrian. *Vision Res* 41, 2487-2502.
- Blakeslee, B., McCourt, M.E., 2004. A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Res* 44, 2483-2503.
- Blakeslee, B., Pasieka, W., McCourt, M.E., 2005. Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Res* 45, 607-615.
- Blankenburg, F., Ruben, J., Meyer, R., Schwiemann, J., Villringer, A., 2003. Evidence for a rostral-to-caudal somatotopic organization in human primary somatosensory cortex with mirror-reversal in areas 3b and 1. *Cereb Cortex* 13, 987-993.
- Blasdel, G.G., Lund, J.S., Fitzpatrick, D., 1985. Intrinsic connections of macaque striate cortex: axonal projections of cells outside lamina 4C. *J Neurosci* 5, 3350-3369.
- Bonhoeffer, T., Kim, D.S., Malonek, D., Shoham, D., Grinvald, A., 1995. Optical imaging of the layout of functional domains in area 17 and across the area 17/18 border in cat visual cortex. *Eur J Neurosci* 7, 1973-1988.
- Boyaci, H., Fang, F., Murray, S. O., and Kersten, D., 2007. Responses to brightness variations in early human visual cortex. *Current Biology* 17, 989-993.
- Brown, J.W., Braver, T.S. 2005. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* 307, 1118-1121.
- Burt, P. J., Adelson, E. H., 1983. The Laplacian pyramid as a compact image code. *IEEE Transaction on Communications* 31, 532-540. Chadderdon, G.L., Sporns, O., 2006. A large-scale neurocomputational model of task-oriented behavior selection and working memory in the prefrontal cortex. *J. Cogn Neurosci* 18, 242-257. 23
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience* 3, 201-215. Corchs, S., Deco, G., 2002. Large-scale neural model for visual attention: Integration of experimental single-cell and fMRI data. *Cereb Cortex* 12, 339-348.
- 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. *J Neurosci* 26, 3634-3641.
- Cornsweet, 1970. Visual perception, Academic Press, New York. Dakin, S.C., Bex, P.J., 2003. Natural image statistics mediate brightness 'filling in'. *Proc Biol Sci* 270, 2341-2348.

## References

- De Valois, R.L., 1971. Contribution of different lateral geniculate cell types to visual behavior. *Vision Res Suppl* 3, 383-396. De Valois, R.L., Webster, M.A., De Valois, K.K., Lingelbach, B., 1986. Temporal properties of brightness and color induction. *Vision Res* 26, 887-897.
- Deco, G., Rolls, E.T., Horwitz, B., 2004. "What" and "where" in visual working memory: a computational neurodynamical perspective for integrating fMRI and single-neuron data. *J Cogn Neurosci* 16, 683-701.
- Denett, D., 1991. Consciousness explained. Boston: Little Brown. DeYoe, E.A., Carman, G.J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., Miller, D., Neitz J. 1996. Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci U S A* 93, 2382-2386.
- DeYoe, E.A., Van Essen, D.C., 1988. Concurrent processing streams in monkey visual cortex. *Trends Neurosci* 11, 219-226.
- Esposito, F., Aragri, A., Piccoli, T., Tedeschi, G., Goebel, R., Di Salle F., 2009. Distributed analysis of simultaneous EEG-fMRI time-series: modeling and interpretation issues. *Magn Reson Imaging* 27, 1120-1130.
- Fernández, G., de Greiff, A., von Oertzen, J., Reuber, M., Lun, S., Klaver, P., Ruhlmann, J., Reul, J., Elger, C.E., 2001. Language mapping in less than 15 minutes: real-time functional MRI during routine clinical investigation. *Neuroimage* 14, 585-594.
- Formisano, E., Kim, D.S., Di Salle, F., van de Moortele, P.F., Ugurbil, K., Goebel, R., 2003. Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron* 40, 859-869.
- Fitzpatrick, D., Lund, J.S., Blasdel, G.G., 1985. Intrinsic connections of macaque striate cortex: afferent and efferent connections of lamina 4C. *J Neurosci* 5, 3329-3349.
- Friedman, H.S., Zhou, H., von der Heydt, R., 2003. The coding of uniform colour figures in monkey visual cortex. *J Physiol* 548, 593-613.
- Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., Turner, R., 1998. Event-related fMRI: characterizing differential responses. *Neuroimage* 7, 30-40.
- Gerrits, H.J., De Haan, B., Vendrik, A.J., 1966. Experiments with retinal stabilized images. Relations between the observations and neural data. *Vision Res* 6, 427-440.
- Gerrits, H.J., Vendrik, A.J., 1970. Simultaneous contrast, filling-in process and information processing in man's visual system. *Exp Brain Res* 11, 411-430.
- Girard, P., Hupé, J., Bullier, J., 2001. Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. *J Neurophysiol* 85, 1328-1331.
- Goebel, R., 1993. Perceiving complex visual scenes: An oscillator neural network model that integrates selective attention, perceptual organisation, and invariant recognition. In: Giles, C., Hanson, S., Cowan, J. (Eds.), *Advances in neural information processing systems* 5. Morgan Kaufmann, San Diego.

## 4 Brightness Induction Modeling

- Goebel, R., Horwitz, B. (in preparation). Integrating Neuroimaging Data and Large-Scale Neural Network Models in a Common Brain Space.
- Goebel, R., De Weerd, P., 2009. Perceptual filling-in: from experimental data to neural network modeling. In: Gazzaniga, M. (Ed.), *The cognitive neurosciences IV*. MIT Press, Cambridge.
- Gove, A., Grossberg, S., Mingolla, E., 1995. Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neurosci* 12, 1027–1052.
- Grimson, W. E., 1982. A computational theory of visual surface interpolation. *Philos. Trans. R. Soc. London* 298, 395–427.
- Grossberg, S., 1987a. Cortical dynamics of three-dimensional form, color, and brightness perception: I. Monocular theory. *Percept Psychophys* 41, 87–116.
- Grossberg, S., 1987b. Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory. *Percept Psychophys* 41, 117–158.
- Grossberg, S., 2003. Filling-in the forms: Surface and boundary interaction in visual cortex. In: Pessoa, L., De Weerd, P. (Eds.), *Filling-in: From Perceptual Completion to Skill Learning*. University Press, Oxford, pp. 13–37.
- Grossberg, S., Raizada, R., 2000. Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vis Res* 40, 1413–1432.
- Grossberg, S., Todorovic, D., 1988. Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Percept Psychophys* 43, 241–277.
- Heider, B., Meskenaite, V., Peterhans, E., 2000. Anatomy and physiology of a neural mechanism defining depth order and contrast polarity at illusory contours. *Eur J Neurosci* 12, 4117–4130.
- Heitger, F., von der Heydt, R., Peterhans, E., Rosenthaler, L., Kübler, O., 1998. Simulation of neural contour mechanisms: representing anomalous contours. *Image and Vision Computing* 16, 409–423.
- Herd, S.A., Banich, M.T., O'Reilly, R.C., 2006. Neural mechanisms of cognitive control: an integrative model of stroop task performance and fMRI data. *J Cogn Neurosci*, 18, 22–32.
- Hinds, O., Polimeni, J.R., Rajendran, N., Balasubramanian, M., Amunts, K., Zilles, K., Schwartz, E.L., Fischl, B., Triantafyllou, C., 2009. Locating the functional and anatomical boundaries of human primary visual cortex. *Neuroimage* 46, 915–922.
- Huang, X., Paradiso, M.A., 2008. V1 response timing and surface filling-in. *J Neurophysiol* 100, 539–547.
- Hubel, D.H., Livingstone, M.S., 1987. Segregation of form, color, and stereopsis in primate area 18. *J Neurosci* 7, 3378–3415.
- Hubel, D.H., Wiesel, T.N., 1959. Receptive fields of single neurones in the cat's striate cortex. *J Physiol* 148, 574–591.
- Hubel, D.H., Wiesel, T.N., 1962. Receptive fields, binocular interaction and functional

## References

- architecture in the cat's visual cortex. *J Physiol* 160, 106-154.
- Hughes, H.C., Nozawa, G., Kitterle, F., 1996. Global precedence, spatial frequency channels, and the statistics of natural images. *J Cogn Neurosci*, 8, 197-230.
- Hung, C.P., Ramsden, B.M., Chen L.M., Roe, A.W., 2001. Building surfaces from borders in Areas 17 and 18 of the cat. *Vision Res* 41, 1389-1407.
- Hung, C.P., Ramsden, B.M., Roe, A.W., 2007. A functional circuitry for edge-induced brightness perception. *Nat Neurosci* 10, 1185-1190.
- Husain, F.T., Tagamets, M.A., Fromm, S.J., Braun, A.R., Horwitz, B., 2004. Relating neuronal dynamics for auditory object processing to neuroimaging activity. *NeuroImage* 21, 1701-1720.
- Jehee, J.F., Lamme, V.A., Roelfsema, P.R. 2007. Boundary assignment in a recurrent network architecture. *Vision Res* 47, 1153-1165.
- Jones, J.P., Palmer, L.A., 1987. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J Neurophysiol* 58, 1233-1258.
- Kanwisher, N., McDermott, J., & Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17, 4302-4311.
- Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R., Ungerleider, L.G., 1999. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751-761.
- Kay, K.N., Naselaris, T., Prenger, R.J., Gallant, J.L., 2008. Identifying natural images from human brain activity. *Nature* 452, 352-355.
- Kayama, Y., Riso, R.R., Bartlett, J.R., Doty, R.W., 1979. Luxotonic responses of units in macaque striate cortex. *J Neurophysiol* 42, 1495-1517.
- Kinoshita, M., Komatsu, H., 2001. Neural representation of the luminance and brightness of a uniform surface in the macaque primary visual cortex. *J Neurophysiol* 86, 2559-2570.
- Kriegeskorte, N., Goebel, R., 2001. An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical mr volumes. *NeuroImage* 14, 329-346.
- Lamme, V. A. F., 1995. The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 10, 649-669.
- Landisman, C.E., Ts'o, D.Y., 2002. Color processing in macaque striate cortex: relationships to ocular dominance, cytochrome oxidase, and orientation. *J Neurophysiol* 87, 3126-3137.
- 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.
- Lu, H.D., Roe, A.W., 2008. Functional organization of color domains in V1 and V2 of macaque monkey revealed by optical imaging. *Cereb Cortex* 18, 516-533.
- Maguire, W.M., Baizer, J.S., 1982. Luminance coding of briefly presented stimuli in area 17 of the rhesus monkey. *J Neurophysiol* 47, 128-137.

#### *4 Brightness Induction Modeling*

- Maier, A., Wilke, M., Aura, C., Zhu, C., Ye, F.Q., Leopold, D.A., 2008. Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nat Neurosci* 11, 1193-1200.
- McCourt, M.E., 1982. A spatial frequency dependent grating-induction effect. *Vision Res* 22, 119-134.
- Mendola, J.D., Conner, I.P., Sharma, S., Bahekar, A., Lemieux, S., 2006 . fMRI Measures of perceptual filling-in in the human visual cortex. *J Cogn Neurosci* 18, 363-375.
- Neumann, H., Pessoa, L., Hansen, T., 2001. Visual filling-in for computing perceptual surface properties. *Biol Cybern* 85, 355-369.
- Neumann, H., Pessoa, L., Mingolla, E., 1998. A neural network architecture of brightness perception: Non-linear contrast detection and geometry-driven diffusion. *Image and Vision Comp* 16, 423-446.
- O'Reilly, R.C., 2006. Biologically Based Computational Models of High-Level Cognition. *Science* 314, 91-94.
- Otazu, X., Vanrell, M., Párraga, C., 2008. Multiresolution wavelet framework models brightness induction effects. *Vision Res* 48, 733-751.
- Paradiso, M.A., Nakayama, K., 1991. Brightness perception and filling-in. *Vision Res* 31, 1221-1236.
- Pereverzeva, M., Murray, S.O., 2008. Neural activity in human V1 correlates with dynamic lightness induction. *J Vis* 8, 1-10.
- Pereverzeva, M., Murray, S.O., 2009. Distant background information strongly affects lightness perception in dynamic displays. *J Vis* 9, 11-10.
- Perna, A., Tosetti, M., Montanaro, D., Morrone, M.C., 2005. Neuronal mechanisms for illusory brightness perception in humans. *Neuron* 47, 645-651.
- Pessoa, L., De Weerd, P. (Eds.), 2003. *Filling-in: From Perceptual Completion to Skill Learning*. Oxford University Press, Oxford.
- Peterhans, E., von der Heydt, R., 1989. Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *J Neurosci* 9, 1749-1763.
- Purves, D., Shimpi, A., Lotto, R.B., 1999. An empirical explanation of the cornsweet effect. *J Neurosci* 19, 8542-8551.
- Raijada, R., Grossberg, S., 2001. Context-sensitive bindings by the laminar circuits of V1 and V2: A unified model of perceptual grouping, attention, and orientation contrast. *Visual Cogn* 8, 431-466.
- Rodieck, R.W., 1965. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Res* 5, 583-601.
- Roe, A.W., Lu, H.D., Hung, C.P., 2005. Cortical processing of a brightness illusion. *Proc Natl Acad Sci U S A* 102, 3869-3874.
- Roebroeck, A., Formisano, E., Goebel, R., 2005. Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage* 25, 230-242.
- Roelfsema, P. R., Lamme, V. A., Spekreijse, H., Bosch, H., 2002. Figure-ground segregation in a recurrent network architecture. *J Cogn Neurosci* 14, 525-537.

## References

- Rossi, A.F., Paradiso, M.A., 1999. Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *J Neurosci* 19, 6145-6156. Rossi, A.F., Rittenhouse, C.D., Paradiso, M.A., 1996. The representation of brightness in primary visual cortex. *Science* 273, 1104-1107.
- Sasaki, Y., Watanabe, T., 2004. The primary visual cortex fills in color. *Proc Natl Acad Sci U S A* 101, 18251-18256.
- Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889-893.
- Shoham, D., Hubener, M., Schulze, S., Grinvald, A., Bonhoeffer, T., 1997. Spatio-temporal frequency domains and their relation to cytochrome oxidase staining in cat visual cortex. *Nature* 385, 529-533.
- Stromeyer, C.F., 3rd, Kronauer, R.E., Madsen, J.C., 1984. Adaptive processes controlling sensitivity of short-wave cone pathways to different spatial frequencies. *Vision Res* 24, 827-834.
- Tagamets, M.A., Horwitz, B., 1998. Integrating electrophysiological and anatomical experimental data to create a large-scale model that simulates a delayed match-to-sample human brain imaging study. *Cereb Cortex* 8, 310-320.
- Tehovnik, E.J., Sommer, M.A., Chou, I.H., Slocum, W.M., Schiller, P.H., 2000. Eye fields in the frontal lobes of primates. *Brain Res Brain Res Rev* 32, 413-448.
- Tootell, R.B., Hadjikhani, N., Hall, E.K., Marrett, S., Vanduffel, W., Vaughan, J.T., Dale, A.M., 1998. The retinotopy of visual spatial attention. *Neuron* 21, 1409-1422.
- Troxler, D., 1804. Ueber das Verschwindern gegebener Gegenstände innerhalb unsers Gesichtskreises. *Ophthalmologisches Bibliothek*. J. Schmidt. Jena, Springer, 431-573.
- Ts'o, D.Y., Roe, A.W., Gilbert, C.D., 2001. A hierarchy of the functional organization for color, form and disparity in primate visual area V2. *Vision Res* 41, 1333-1349.
- von der Heydt, R., Peterhans, E., 1989. Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *J Neurosci* 9, 1731-1748. von der Heydt, R., Peterhans, E., Baumgartner, G., 1984. Illusory contours and cortical neuron responses. *Science* 224, 1260-1262.
- Wang, Y., Xiao, Y., Felleman, D.J., 2007. V2 thin stripes contain spatially organized representations of achromatic luminance change. *Cereb Cortex* 17, 116-129.
- Wachtler, T., Sejnowski, T.J., Albright, T.D., 2003. Representation of color stimuli in awake macaque primary visual cortex. *Neuron* 37, 681-691.
- White, M., 1979. A new effect of pattern on perceived lightness. *Perception* 8, 413-416.
- Yabuta, N.H., Callaway, E.M., 1998. Cytochrome-oxidase blobs and intrinsic horizontal connections of layer 2/3 pyramidal neurons in primate V1. *Vis Neurosci* 15, 1007-1027.
- Yacoub, E., Harel, N., Ugurbil, K., 2008. High-field fMRI unveils orientation columns in humans. *Proc Natl Acad Sci U S A* 105, 10607-10612.



## **5 A spatio-temporal correlate of texture filling-in in human V2**

based on: Jans, B., Been, M.M., Goebel, R. & De Weerd, P., A spatio-temporal correlate of texture filling-in in human V2. (submitted)

## Abstract

During maintained fixation away from a figure, the figure becomes perceptually filled in by the surrounding background. For brightness and color, some studies, but not all, have demonstrated a signal within the extent of the figure's retinotopic projection in early visual areas correlating with perceptual filling-in. Here, we investigated the neural correlates of a more complex surface feature, namely texture. While human participants fixated a point away from a grey figure surrounded by dynamic texture, we observed fMRI signal within the confines of the figure's retinotopic projection in V2 that correlated with filling-in ratings collected during scanning. Moreover, perceptual filling-in and its correlate in V2 were strongest during accurate fixation, and while the fluctuating signal elicited by the background texture was highest. Hence, we have demonstrated for the first time a link between activity increases and perceptual filling-in of texture in early human visual cortex (in V2, but not V1), thereby confirming and extending earlier observations in a neurophysiological study in monkeys. The data point to the importance of boundary adaptation as a permissive factor for filling-in, to a competitive interaction between figure and background, and to a possible contribution of interpolation processes in V2 to the visual awareness of surfaces.

## Introduction

Perceptual filling-in refers to the visual fading or the disappearance of an object from awareness during stabilization of the object on the retina (Martinez-Conde et al., 2004). We used a paradigm already used by Troxler (Troxler, 1804) in the late 18<sup>th</sup> century, in which observers fixate their eyes on a point away from a figure for several seconds, after which the figure fades and becomes perceptually replaced by the background (Krauskopf, 1963; Gerrits et al., 1966). Filling-in takes place in various stimulus domains including luminance (Roe et al., 2005; Cornelissen et al., 2006; Mendola et al., 2006; Pereverzeva and Murray, 2008), color (von der Heydt et al., 2003; Sasaki and Watanabe, 2004) and texture (De Weerd et al., 1995; Weil et al., 2007; Weil et al., 2008; Weil and Rees, 2010), and has been a favored paradigm to study mechanisms of surface perception.

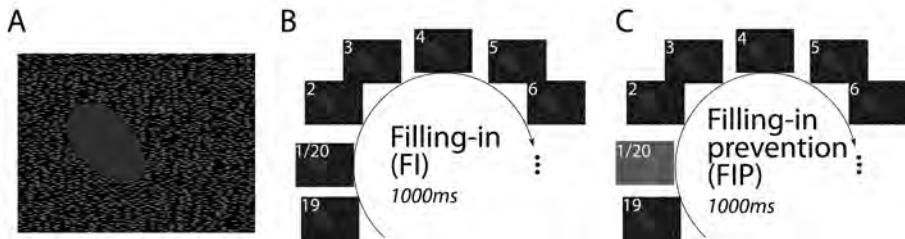
It is well established that the accuracy of fixation determines the delay before filling-in takes place (Troncoso et al., 2008). This finding agrees with proposals in which the spread of surface-related activity is contained

within retinotopically appropriate boundaries by an inhibitory signal from boundary detectors (Walls, 1954; Gerrits et al., 1966; Gerrits and Vendrik, 1970; Grossberg, 1987a, b). According to these proposals, accurate fixation increases the rate of adaptation in the boundary detectors, leading to faster disinhibition, and a spread of background-related activity into the prior retinotopic territory of the figure (interpolation model). The data suggest that the time course of adaptation is slow (lasting seconds), and that once boundary representations fail, spread takes place quickly (on a millisecond time scale) (Paradiso and Nakayama, 1991; Arrington, 1994; De Weerd et al., 1998; Angelucci and Bressloff, 2006; De Weerd, 2006; Huang and Paradiso, 2008). In other models of filling-in, neural signals related to surface perception are thought to result from the properties of linear filters in the visual system (Grimson, 1982; Ohzawa et al., 1985; Grossberg, 1987a, b; Grossberg and Todorovic, 1988; Blakeslee et al., 2005).

Despite compelling behavioral data and computational models, the neural processes underlying surface perception remain poorly understood. Even on the presence of increased activity within the confines of the retinotopic projection of a filled-in surface, which would be expected if spread or interpolation played a role, data are conflicting. Using a variety of paradigms, a number of studies reported positive correlations between neural activity and perceived changes in surface brightness or color (De Weerd et al., 1995; Sasaki and Watanabe, 2004; Meng et al., 2005; Roe et al., 2005; Pereverzeva and Murray, 2008), but other studies reported an absence of activation changes (von der Heydt et al., 2003; Cornelissen et al., 2006), or deactivations (Mendola et al., 2006; Weil et al., 2007; Weil and Rees, 2010). The present study aimed to identify neural processes within the projection of a large grey figure related to perceptual filling-in by a background of dynamic texture, in human participants.

In the only preceding study that investigated neural correlates of texture filling-in using comparable stimuli (De Weerd et al., 1995), monkeys fixated a point away from a grey square surrounded by dynamic texture, and V2 and V3 neurons with RFs contained in the grey figure showed strong activity increases that coincided in time with filling-in reports of humans exposed to the same stimulus conditions. So far, this finding has not been reported in humans. Here, we measured fMRI activity in the retinotopic projection of a grey figure surrounded by equiluminant dynamic texture in humans (Figure 1A) while they reported the strength of the illusion.

## 5 Texture Filling-in



**Figure 1.** Stimulus design and experimental conditions A) Representative stimulus frame: Each frame was composed of randomly positioned horizontal line elements on a black background which served as a background for a grey equi-luminant figure. The center of gravity of the figure was presented at 7° eccentricity (see Experimental stimuli in Experimental procedures section). Participants were instructed to fixate on the small red square. B) Filling-in (FI) condition: Stimulus frames with randomly positioned line elements were refreshed at 20 frames per second, making the texture background dynamic. When fixating the red square participants perceived PFI. C) Filling-in prevention (FIP) control condition: To prevent the filling-in percept from occurring the filling-in the first texture frame of the stimulus movie (played at 20 Hz) was dropped. This has been shown previously to prevent filling-in [ref]. Participants were instructed to perform a demanding hue discrimination task at fixation in order to promote strong fixation.

## Methods

Five healthy and naïve participants (1 male and 4 female, mean age 28 years) with normal or corrected to normal vision gave written informed consent before starting the fMRI experiment, and were debriefed following completion of data acquisition.

### Retinotopic mapping and localizer stimuli

Following standard retinotopic mapping (Engel et al., 1997; Engel et al., 1994; Sereno et al., 1995), an localizing the centre of the figure with a stationary circular checkerboard, an additional set of localizer stimuli was used to determine a more fine grained map of the figure region to make precise spatial inferences possible. An expanding, high contrast green-red checkerboard ring in the same shape as the figure was used. The ring expanded continuously from center to its maximum extent in 40s, while remaining centered at the point of gravity of the figure region (Figure 3B). This resulted in an accurate eccentricity representation and demarcation of the figure region and its surround (Figure 3C). Experimental Stimuli Filling-in stimuli consisted of a grey figure and a texture surround. The texture in all stimuli used was made up of horizontal line segments ( $0.8^\circ$  by  $0.1^\circ$ , posi-

tion 0.5° apart on average, Figure 1A), which randomly changed position on each frame update (Figure 1B). Frames were updated every 50ms resulting in an average luminance of 5.46 cd/m<sup>2</sup> over time. The grey figure was equi-luminant with the background, and its center of gravity was at 7 degrees eccentricity relative to the fixation point (small red square 0.2° by 0.2°). To have the middle of the figure's projection on retinotopic cortex roughly equally far from surrounding boundary representations, we used an egg-like figure derived from an inverse visual field mapping function (Sereno et al., 1995). This approach was devised to maximally protect the center of the figure's projection from fMRI signal spread unrelated to perception. Participants reported filling-in in this stimulus filling-in condition (FI), except when once every 20 frames the texture was replaced by equi-luminance grey (filling-in prevention, FIP, Figure 1C), which counteracted the filling-in process. In the FIP condition, participants performed a demanding hue detection task at fixation. In addition, a dynamic full field texture stimulus was presented, of which the elicited fMRI activity served as a reference against which to compare fMRI activity in FI and FIP conditions.

### **Experimental design and participant instructions**

We used a block design in which 8 experimental block (30s) were interspersed with rest periods (30s) during each run. A total of 9 experimental runs (1 full field texture run, 4 filling-in runs, and 4 filling-in prevention runs) and 5 localizer runs (2 retinotopy runs, 1 full field texture run, 3 figure eccentricity localizer runs) were presented to the participants, while they maintained fixation. To counteract participants' fatigue, we divided the runs over 2 experimental sessions. In the FI condition, participants were instructed to fixate the fixation point and covertly attend to the figure region. They were asked to report the extent to which the texture surround entered the figure region. This filling-in rating (FIR) was carried out by moving a joystick lever from the center (not filled-in) forward to the maximum extent of the throttle (completely filled-in). Participants were asked to map the extent of intrusion of texture into the figure by the extent of forward movement of the joystick, with maximal forward movement signaling that the figure was completely and strongly filled-in. During the FIP condition, participants fixated the fixation point and performed a hue detection task at fixation, in which participants had to detect small and brief deviations from red occurring unpredictably at an average frequency of 0.5 Hz. In the texture condition, participants passively fixated the fixation point (red square).

## Apparatus

A Siemens 3T Allegra MRI scanner (Siemens Erlangen) was used to collect both anatomical and functional images. Anatomical T1-weighted (MP-RAGE; TR: 2250ms; matrix size,  $256 \times 256$ ; 192 slices; in-plane resolution,  $1 \text{ mm}^2$ ) and functional echo planar imaging (EPI) scans (28 gapless slices, TR/TE, 2,000/30ms; flip angle 90°; square in plane matrix size of  $128^2$ ; resulting in isotropic  $2\text{mm}^3$  voxels) were acquired of two participants on a 3T Siemens Allegra head scanner. Because we chose to maximize spatial resolution, the block of functional slices was positioned centered around the calcarine sulcus, covering both dorsal and ventral V1 and V2. Stimuli were presented to the participants via the LCD stereo displays in the Visuastim Digital (Resonance Technology, Northridge USA) with a built-in Arrington Viewpoint eyetracking system (Arrington Research, Scottsdale USA) with a refresh/sample rate of 60 Hz.

Analog ratings of the percept experienced by the participants were collected with a joystick fiber optic response pad (Current Designs, Philadelphia USA). Calibrated rating data were stored at over 1000 Hz, post processing steps included drift correction for each trial and temporal resampling to the scanning repetition time (TR = 2000ms).

## Functional data preprocessing

Functional localizer and retinotopic data were pre-processed in Brainvoyager QX (Goebel, 2010; Goebel et al., 2006) including slice time correction, 3D correction for head volume movements and temporal filtering, which included linear trend removal and a high-pass filter of 3 cycles per time course. The functional data were resampled and co-registered to the Talairach transformed anatomical data (Goebel et al., 2006). Anatomical data were corrected for field inhomogeneity, and volumetrically normalized into Talairach space. A brightness cut-off method was used to identify the white-grey matter border where a mesh was modelled to reconstruct the cortical sheet. The meshes were subsequently inflated, flattened and corrected for spatial distortions. To perform a surface based functional analysis, the volumetric functional data were sampled along the grey-white matter boundary and transformed into a 2D coordinate system with Brainvoyager QX.

## **Functional and psychophysical data analysis**

The localizer data obtained from the continuously expanding rings resulted in a slow wave of activity which was analyzed using a cross correlation analysis with 20 lags. This revealed a detailed mapping of the figure region (in analogy with standard retinotopic mapping).

Based on the stimulation protocol, the experimental blocks were transformed to percent signal change relative to the time points preceding each block (epoch based PSC transform). Condition wise averaging of all blocks across participants resulted in event related averages. The event related averages were split into an early (6-12s) and a late (12-18s) part. Data from the early and late part were averaged for each condition and submitted to ANOVA; Post-hoc comparisons were bonferroni corrected.

In conjunction with retinotopy data and localizer data, cross sectional paths across figure representations were defined in V1v and V2v for all participants separately. Along the cross sectional paths, landmarks indicating the edges and centre of the figure region were defined. By using these landmarks as constraints, activity patterns along the cortical paths could be aligned and averaged without loss in data detail due to variations in path length and landmarks among participants. To link fMRI activity along the cross sectional paths with FIRs, we down sampled the continuous ratings to the time resolution of the 2s TR, and subdivided the ratings into 5 equal bins. By dividing the number of events in each bin by the total number, we computed the FIR in each bin (Figure 2A). We then selected the cross sectional data patterns for fMRI time points corresponding to the highest FIR bin in all participants. Based on this selection of the data, we computed the multi participant cross sections of the figure region for V1v and V2v in filling-in (FI) and filling-in prevention (FIP) conditions. Cross sectional data from the centre region were submitted to an ANOVA with condition as a fixed factor and participant as a random variable. From the cross sectional data, a metric summarizing the amount of overlap between the FI and FIP condition was computed by calculating the surface between the FI and FIP curve (Net Filling-in Index NFI).

The NFI was then computed also for the other 4 rating bins, and the linear relationship between the NFI and FIR across all bins was quantified by a Pearson correlation coefficient. In addition, texture activity levels were obtained from full field texture stimulus, and activity elicited by the texture was correlated with the FIR.

## 5 Texture Filling-in

### Eye movement data

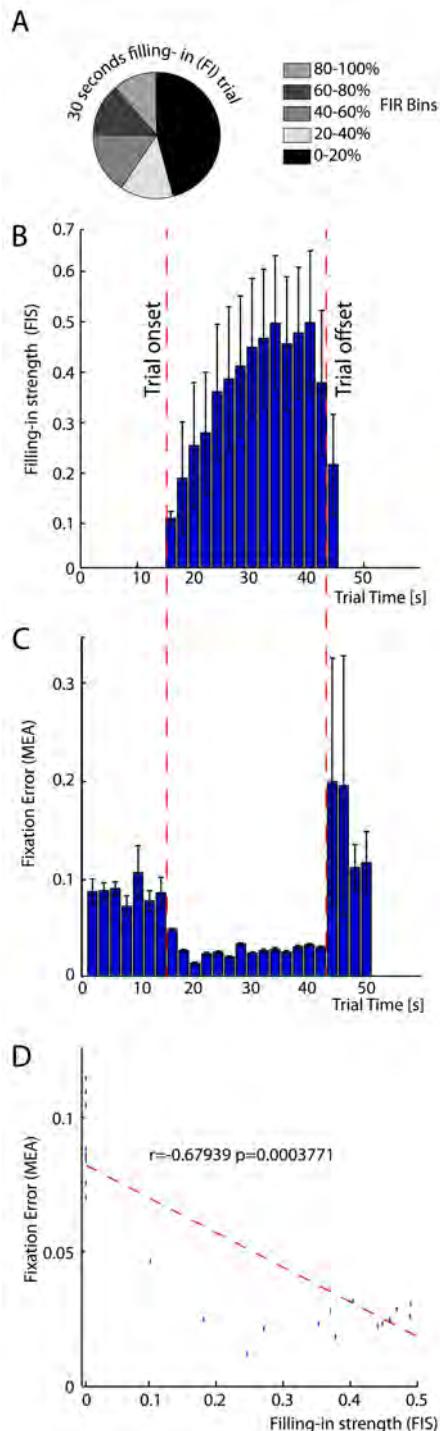
Raw eye movement data were binned with bin duration equal to scanning repetition time (TR = 2000ms). The bin offset was chosen to occur at the onset of the image acquisition. The distribution of gaze points within the duration of a TR was fitted with an ellipse of minimal area (MEA) (Nachmias, 1959; Sun and Freund, 2004), which was used as an index of fixation error.

## Results and discussion

This experimental design permits directly linking perceptual filling-in and its neural correlate, which is important because the link between both reported earlier (De Weerd et al., 1995) was indirect and based on an across-species comparison.

Participants (N=5) were exposed to 30s periods of stimulation with a dynamic texture stimulus that either permitted filling-in (FI condition, Figure 1B), or prevented filling-in (FIP condition, Figure 1C). In the FIP condition, filling-in was prevented by replacing single frames with equiluminant grey each sec [16]. In the FI condition, participants continuously monitored and reported their impression of filling-in using a joystick, and generated a continuous Filling In Rating (FIR). To facilitate the comparison of FIRs with fMRI data, the FIRs were downsampled to the resolution of the TR (2000ms), and reduced to a relative frequency distribution over 5 equal bins representing the extent of joystick movement (0-20, 20-40, 40-60, 60-80, 80-100%), with the highest bin corresponding to (near) complete filling-in (for details see Functional and psychophysical data analysis section in Experimental procedures). Figure 2A shows that in 11% of the TRs, complete filling-in was perceived. In addition, Figure 2B shows that on average, participants perceived the strongest filling-in (FIS, filling-in strength) near the second half of the trial. This is in line with the idea that boundary adaptation – a process which is time consuming – takes place before perceptual filling-in can be perceived.

Throughout the entire experiment fixation behavior was recorded. The eye movement data (Figure 2C) show that fixation accuracy was high during the filling-in trials. Moreover, in line with previous studies (Martinez-Conde et al., 2004; Martinez-Conde et al., 2006), we found that good fixation accuracy facilitated perceptual filling-in (Figure 2D). Accuracy was quantified as the Minimum Ellipsoid Area (MEA) encompassing a cluster of gaze points within every TR. Because the MEA reflects the magnitude of fix-



## 5 Texture Filling-in

### Figure 2. Stimulus ratings and eye movements

- A) Frequency of FIR binned in 5 equal intervals. The histogram (pie chart) shows average relative frequency per bin averaged over all 160 30s filling-in trials across all participants.
- B) Average filling-in strength (FIS) increases over time. From each participant ( $N=5$ ) we collected 32 filling-in trials with a duration of 30 seconds. The FIRs during these trials for 5 participants, a total of 160 trials, were collapsed per TR resulting in the average FIR. A majority of the time participants did not perceive a strong filling-in percept, therefore the FIS distribution was skewed. This was due to the choice of a very large figure, which was necessary to overcome effects of fMRI signal spread (see text). The error bars on the bars represent semi- inter quartile ranges, a measure of dispersion robust to skewness.
- C) Participants fixation errors as a function of trial time. Fixation errors were quantified as minimum ellipsoid area (MEA) of gaze points around fixation, from 160 trials in 5 participants (details in text). Between the trial onset and offset MEA magnitude was minimal, indicating very accurate fixation. Error bars show semi- interquartile ranges, note the large error bars at the trial offset when participants relax their eyes.
- D) High fixation accuracy predicts average FIS. The strong negative correlation between high fixation accuracy (small MEA) and FIS results in a significantly negative correlation ( $r=-0.67939$ ;  $p= 0.0003$ ).

ation errors in squared visual degrees, a low MEA corresponds to a high FIS which was revealed by a negative correlation between MEA and FIS ( $r=-0.65$ ;  $p=0.003$ ,  $N=5$ ).

To identify an fMRI correlate of filling-in, signal during FI and FIP trials (blocks) was compared. However, large fMRI signal spread into the representation of the figure was expected in both FI and FIP conditions, due to the low spatial resolution of fMRI (Cornelissen et al., 2006; Engel et al., 1994). Therefore, any neural correlate of filling-in in the figure's representation might be undetectable against the large effects of fMRI signal spread. Hence, we chose a size for the grey shape ( $34^{\circ}$ ) that based on a prior study using shapes at the same eccentricity is close to the largest size that would still permit filling-in (about  $36^{\circ}$  (De Weerd et al., 1998)), and then developed a set of localizers to precisely determine the cortical projection of the center of the figure in retinotopic visual areas V1 and V2 (Figure 3). The central representation of the grey shape is farthest removed from the effect of intrinsic spread of fMRI signal, and we expected that in the

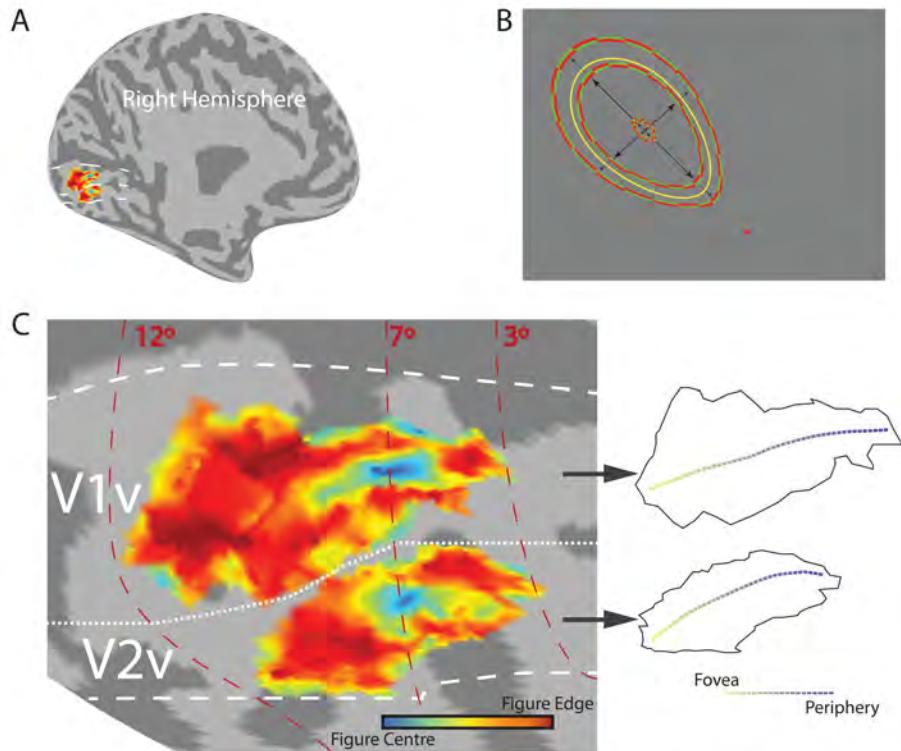
central representations of V1 and V2, the likelihood of finding a correlate of perceptual filling-in would be maximized.

Using event-related averaging (based on 160, 30s stimulus presentations; 32 events x 5 participants), we found strong activation inside the overall figure representation (defined by the eccentricity localizer of the figure) in the full-field texture condition (4.63% change against baseline), and in the FIP condition (3.75%). This was true in V1 (Figure 4A) and V2 (Figure 4B). Note that activity in the FIP condition was significantly but only slightly smaller than with a full field texture, witnessing of the enormous signal spread from the texture background into the figure representation. Despite this signal spread, we observed a dynamic increase in fMRI signal with time in the FI condition. To test this effect, we carried out a 3-way ANOVA including the factors Condition (FI, FIP, Tex), area (V1, V2), and Time Window (1st vs 2nd half of trial). We observed a two way interaction between area and condition ( $F[2]=15.59$ ,  $p>0.001$ ) and between time window and condition ( $F[2]=11.75$ ,  $p>0.001$ ). We then performed bonferroni-corrected ( $\alpha/18$ ; area(3) x timewindow(2) x condition(3)) post hoc tests separating visual areas, and early and late intervals.

In V1, activity in the early time interval did not differ significantly between FI and FIP conditions (Figure 4A, early interval, blue vs. green bars), and this remained true in the late interval. In V2, activity in the early time interval did not differ between FI and FIP conditions (Figure 4B, early interval, blue vs. green bars), but in the late interval activity in the FI condition significantly differed from that in the FIP condition (Figure 4B, late interval, blue vs. green bars). Moreover, activity in the FI and full texture conditions became indistinguishable in the late time window (Figure 4B, late interval, blue vs. red bars).

We then aimed to correlate fMRI signal in the center of the figure's representation in V2 with filling-in reports as they fluctuated during the length of single trials. To that aim, we used the FIRs downsampled to the resolution of the TR (Figure 2A), and reduced to a relative frequency distribution over 5 equal bins (0-20, 20-40, 40-60, 60-80, 80-100%). The events in each of these bins served as a regressor for computing a profile of Beta-values along a cross section (cortical distance axis in Figure 3C) of the figure representation for the FI and FIP conditions. Figure 4C shows these profiles for a regressor based on events belonging to the bin of strongest FIRs. Along the cortical distance axis in V2 we found a highly significant signal increase of the FI condition as compared to the FIP condition in the center of the figure region in V2 ( $F(3,4) = 10,52$ ;  $p = 1.09 \times 10^{-6}$ ).

## 5 Texture Filling-in



**Figure 3.** fMRI localizer data

A) Retinotopic localization of the figure region: Mapping results projected on a segmented right hemisphere.

B) Mapping stimuli: The extent of the grey figure was mapped by presenting green/red reversing oval outlines which slowly moved from the center towards the edge. The yellow outline shows the extent of the grey figure, the mapping included a part of the texture surround. C Enlarged results of mapping the figure region in a single participant. The localization procedure results in a systematic mapping of distance from center (blue) to edge (red) of the figure's cortical projection in V1v and V2v (see supplementary materials), as the stimulus was presented in left upper quadrant and represented in the lower bank of the calcarine sulcus. Red dash lines are iso-eccentricity lines based on eccentricity mapping, white dash lines represent the horizontal meridians, white dotted lines the vertical meridian. The statistical map outlines on the right show the cross section path through the center of the figure representation along a line of equal polar angle from the fovea (yellow) to the periphery (blue) in V1v and V2v. The paths were defined individually for each hemisphere and visual area, and their position visualized in statistical map outlines.

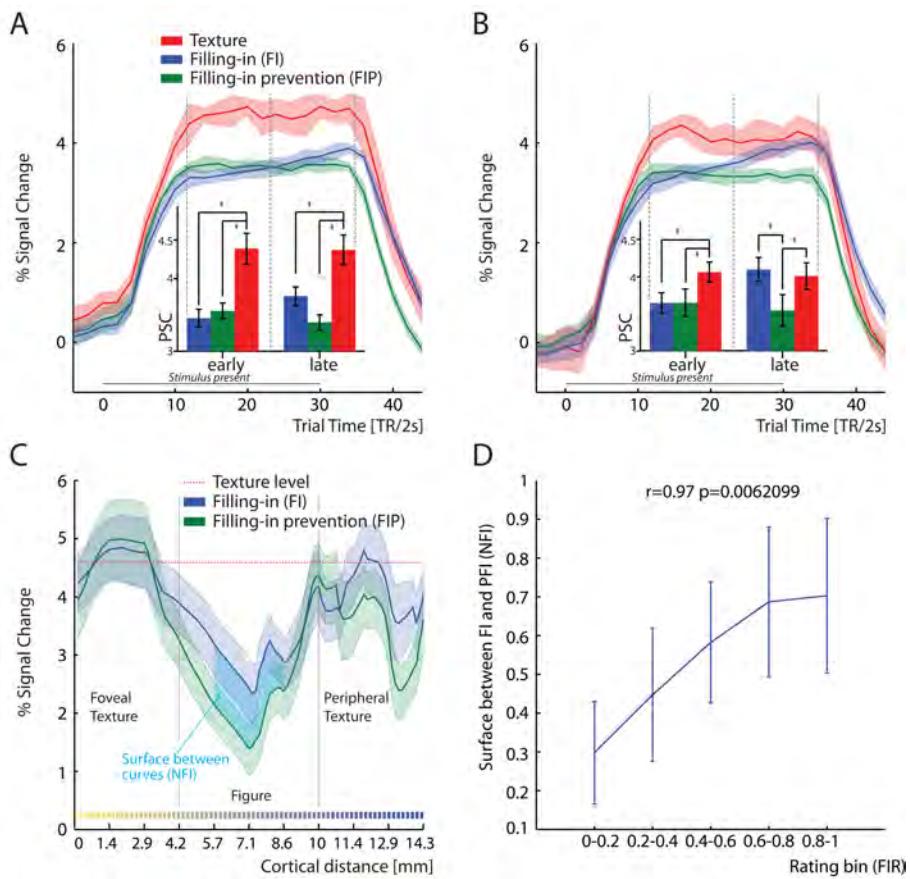
## *Results and discussion*

In agreement with neurophysiologic findings (De Weerd et al., 1995) this result was not found in V1 ( $F(3,4) = 1,12$ ;  $p = 0,3405$ ). The difference between the activity profiles in FI and FIP conditions can be quantified by the surface contained between the two curves (Net Filling-in Index (NFI), see Figure 4C). Figure 4D reveals a high correlation between the FIR and their corresponding NFI, ( $r = 0.96$ ,  $p = 0.0094$ ). This shows that the area between the FI and FIP curve in the center of the figure region (NFI) is directly related to the strength of the filling-in process as perceived by the participant (FIR).

The data provide support for the presence of activity within the confines of the retinotopic projection that is relevant for surface perception. In agreement with neurophysiological data (De Weerd et al., 1995), this activity appears to be present in extrastriate cortex, but not in V1. Further, the increasing activity found in human V2 (Figure 3B) is remarkably similar to that reported in monkey extrastriate areas V2 and V3. However, whereas the activity increases in neurons with RFs inside the grey figure region were recorded while the monkey ignored the stimuli, the present experiment permitted for the first time to directly correlate activity modulations inside the figure representations with strength of perceptual filling-in.

Furthermore, we found that filling-in was more probable during accurate fixation, which agrees with the idea that boundary stabilization and disinhibition leads to surface feature spread (Walls, 1954; Gerrits et al., 1966; Gerrits and Vendrik, 1970; Grossberg, 2003). We also found that higher activity in the texture region immediately surrounding the figure representation during filling-in episodes (2.5 mm on cortex extending into the background representation from the figure edge representation) correlated with the NFI based on the strongest FIR bin ( $r = 0.906$ ,  $p = 0.034$ ,  $n=5$ ). This suggests that while boundary adaptation progresses, perceptual filling-in is stronger and more likely when the texture surrounding the figure happens to elicit stronger activity. This indicates a competitive relationship between background texture representations and boundary (or figure) representations (De Weerd et al., 1998). This competitive relationship is in line with prior findings indicating that the larger surface (seen as the background) will fill in the smaller part of a stimulus (seen as the figure), with probabilities predicted by their relative sizes (De Weerd et al., 1998; Hsieh and Tse, 2006). Recent reports of 'feature mixing' during maintained fixation in displays made of equal size 'background' and 'figure' areas suggest that also surface features themselves may compete during filling-in (Hsieh and Tse, 2009, 2010).

## 5 Texture Filling-in



**Figure 4.** fMRI results of the filling-in paradigm

A) fMRI activity during filling-in and control conditions of the entire figure region in V1. Event related average ( $N=5$ , error bar = 1 SEM) shows V1 fMRI signal as a function of time in the FI (green), FIP (blue) and texture (red) stimulus conditions. The bar graph below the time series show the average activity level of the event related averages for an early and late interval (indicated by grey vertical lines). Error bars on the bars are 95% confidence intervals, linked bars marked with an asterisk differ significantly (Bonferroni corrected). When comparing activity levels between time intervals, the only significant increase was found in the FI condition in V2 (not shown) ( $F(2) = 1.35$ ,  $p = 0.001$ ).

B) fMRI activity during filling-in and control conditions of the entire figure region in V2 (conventions as in A).

C) Activity profiles along an anatomical path across the figure representation in V2 for filling-in (FI) and Filling-in prevention (FIP) conditions. Activity along the cortical path (see Figure 3C) in FI (blue line) and FIP (green line) conditions is shown from more foveal to more peripheral locations. For reference, to the average activity level for the full field texture stimulus is shown (thin red dashed line). Along and above the abscissa a color gradient is shown corresponding to the cortical sampling path as (Figure 3B) through V2. To outline the localizer data in this figure, grey vertical lines are used to delineate the edges of the figure or the transition from the figure region into the texture. The area used to estimate the fMRI correlate of filling-in is the blue shaded area between the FI an FIP curve (NFI) and extends over 2.9 mm cortex along the line sample.

D) fMRI signal correlates with texture filling-in. The fMRI correlate was defined as the NFI (see Figure 4C), and is correlated with the FIR.

The presence of enhanced activity in surface representations during filling-in, the evidence for a role of boundary adaptation, and the competitive relationship between background and boundary are all elements in our data that are consistent with interpolation models of surface perception. However, in a recent interesting study, Weil et al. (Weil et al., 2007; Weil et al., 2008) concluded that texture filling-in is associated with a decrease in fMRI and MEG signal. Importantly, Weil et al.'s experiments differed in an important aspect from ours: while we used a large static grey figure region, they used a very small flickering target. It is possible that the decrease in activity found in their study reflected adaptation to the flickering target, rather than a correlate of texture filling-in. Our data are in support of that

## *5 Texture Filling-in*

interpretation, and although our data do not support interpolation directly, they confirm that surface texture interpolation can be associated with an excitatory process within the extent of a surfaces' retinotopic projection in early visual areas. The same conclusion is supported by a number of studies in the domains of color and brightness (De Weerd et al., 1995; Sasaki and Watanabe, 2004; Meng et al., 2005; Roe et al., 2005; Pereverzeva and Murray, 2008).

We present the first human evidence linking increases in the levels of fMRI activity in the representation of a physically grey surface with the percept of texture filling-in across that surface. The fMRI data show strong convergence with neurophysiological data from the monkey obtained with similar stimuli. However, the temporal resolution of fMRI does not permit to test directly whether the activity increases correlated with perceptual filling-in reflect a fast neural interpolation process. Further neurophysiological experiments in monkeys may be an effective manner to investigate this open question.

## **References**

- Angelucci, A., and Bressloff, P.C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Prog Brain Res* 154, 93-120.
- Arrington, K.F. (1994). The temporal dynamics of brightness filling-in. *Vision Res* 34, 3371-3387.
- Blakeslee, B., Pasieka, W., and McCourt, M.E. (2005). Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Res* 45, 607-615.
- Cornelissen, F.W., Wade, A.R., Vladusich, T., Dougherty, R.F., and Wandell, B.A. (2006). No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex. *J Neurosci* 26, 3634-3641.
- De Weerd, P. (2006). Perceptual filling-in: More than the eye can see. *Prog Brain Res* 154, 227-245.
- De Weerd, P., Desimone, R., and Ungerleider, L.G. (1998). Perceptual filling-in: a parametric study. *Vision Res* 38, 2721-2734.
- De Weerd, P., Gattass, R., Desimone, R., and Ungerleider, L.G. (1995). Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature* 377, 731-734.
- Engel, S.A., Glover, G.H., and Wandell, B.A. (1997). Retinotopic organization in human

## References

- 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., and Shadlen, M.N. (1994). fMRI of human visual cortex. *Nature* 369, 525.
- Gerrits, H.J., De Haan, B., and Vendrik, A.J. (1966). Experiments with retinal stabilized images. Relations between the observations and neural data. *Vision Res* 6, 427-440.
- Gerrits, H.J., and Vendrik, A.J. (1970). Simultaneous contrast, filling-in process and information processing in man's visual system. *Exp Brain Res* 11, 411-430.
- Goebel, R. (2010). Brainvoyager QX. (Maastricht, Brain Innovation). Goebel, R., Esposito, F., and Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with Brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp* 27, 392-401.
- Grimson, W.E. (1982). A computational theory of visual surface interpolation. *Philos Trans R Soc Lond B Biol Sci* 298, 395-427.
- Grossberg, S. (1987a). Cortical dynamics of three-dimensional form, color, and brightness perception: I. Monocular theory. *Percept Psychophys* 41, 87-116.
- Grossberg, S. (1987b). Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory. *Percept Psychophys* 41, 117-158.
- Grossberg, S. (2003). Filling-in the forms: Surface and boundary interaction in visual cortex. In *Filling-in: From Perceptual Completion to Skill LEarning*, L. Pessoa, and P. De Weerd, eds. (Oxford: University Press), pp. 13-37.
- Grossberg, S., and Todorovic, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: a unified model of classical and recent phenomena. *Percept Psychophys* 43, 241-277.
- Hsieh, P.J., and Tse, P.U. (2006). Illusory color mixing upon perceptual fading and filling-in does not result in 'forbidden colors'. *Vision Res* 46, 2251-2258.
- Hsieh, P.J., and Tse, P.U. (2009). Feature mixing rather than feature replacement during perceptual filling-in. *Vision Res* 49, 439-450.
- Hsieh, P.J., and Tse, P.U. (2010). "Brain-reading" of perceived colors reveals a feature mixing mechanism underlying perceptual filling-in in cortical area V1. *Hum Brain Mapp* 31, 1395-1407.
- Huang, X., and Paradiso, M.A. (2008). V1 response timing and surface filling-in. *J Neurophysiol* 100, 539-547.
- Krauskopf, J. (1963). Effect of retinal image stabilization on the appearance of heterochromatic targets. *J Opt Soc Am* 53, 741-744.
- Martinez-Conde, S., Macknik, S.L., and Hubel, D.H. (2004). The role of fixational eye movements in visual perception. *Nat Rev Neurosci* 5, 229-240.
- Martinez-Conde, S., Macknik, S.L., Troncoso, X.G., and Dyar, T.A. (2006). Microsaccades counteract visual fading during fixation. *Neuron* 49, 297-305.

## 5 Texture Filling-in

- Mendola, J.D., Conner, I.P., Sharma, S., Bahekar, A., and Lemieux, S. (2006). fMRI Measures of perceptual filling-in in the human visual cortex. *J Cogn Neurosci* 18, 363-375.
- Meng, M., Remus, D.A., and Tong, F. (2005). Filling-in of visual phantoms in the human brain. *Nat Neurosci* 8, 1248-1254. Nachmias, J. (1959). Two-dimensional motion of the retinal image during monocular fixation. *J Opt Soc Am* 49, 901-908.
- Ohzawa, I., Sclar, G., and Freeman, R.D. (1985). Contrast gain control in the cat's visual system. *J Neurophysiol* 54, 651-667.
- Paradiso, M.A., and Nakayama, K. (1991). Brightness perception and filling-in. *Vision Res* 31, 1221-1236. Pereverzeva, M., and Murray, S.O. (2008). Neural activity in human V1 correlates with dynamic lightness induction. *J Vis* 8, 8 1-10.
- Roe, A.W., Lu, H.D., and Hung, C.P. (2005). Cortical processing of a brightness illusion. *Proc Natl Acad Sci U S A* 102, 3869-3874.
- Sasaki, Y., and Watanabe, T. (2004). The primary visual cortex fills in color. *Proc Natl Acad Sci U S A* 101, 18251-18256.
- Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889-893.
- Sun, P., and Freund, R.M. (2004). Computation of Minimum-Volume Covering Ellipsoids. *Operations Research* 52, 690-706.
- Troncoso, X.G., Macknik, S.L., and Martinez-Conde, S. (2008). Microsaccades counteract perceptual filling-in. *J Vis* 8, 15 11-19.
- Troxler, I. (1804). Über das Verschwinden gegebener Gegenstände innerhalb unseres Gesichtskreises. *Ophthalmologische Bibliothek* 2, 1-53.
- von der Heydt, R., Friedman, H.S., and Zhou, H. (2003). Searching for the Neural Mechanisms for Color Filling-In. In *Filling-in From Perceptual Completion to Cortical Reorganization*, P. De Weerd, and L. Pessoa, eds. (New York: Oxford University Press).
- Walls, G.L. (1954). The filling-in process. *Am J Optom Arch Am Acad Optom* 31, 329-341.
- Weil, R.S., Kilner, J.M., Haynes, J.D., and Rees, G. (2007). Neural correlates of perceptual filling-in of an artificial scotoma in humans. *Proc Natl Acad Sci U S A* 104, 5211-5216. Weil, R.S., and Rees, G. (2010). A new taxonomy for perceptual filling-in. *Brain Res Rev.*
- Weil, R.S., Watkins, S., and Rees, G. (2008). Neural correlates of perceptual completion of an artificial scotoma in human visual cortex measured using functional MRI. *Neuroimage* 42, 1519-1528.

## **6 Discussion and summary**

## **Summary**

An overwhelming amount of visual processing resources is devoted to the reconstruction of surface boundaries based on fragmentary retinal input. Boundary reconstruction is well-studied, and there is common agreement that neural interpolation processes contribute to this reconstruction process. However, the signal elicited by the interior of homogenous surfaces is small, and computational models as well as empirical studies disagree on the question whether this signal has any relevance for the subjective perception of the interior of surfaces. The studies presented here were focused on the principal, heavily debated question whether there is a signal associated with the perception of a surface's interior within the confines of its retinotopic representation in early visual cortex. To address this question we performed fMRI studies at 3T using two different psychophysical paradigms to manipulate the perceptual aspect of surfaces in the absence of physical surface changes.

Prior to investigating this question, a method for efficient functional spatial alignment was developed to minimize signal loss during averaging due to cross-individual differences. This methodological development was necessary in view of the small signals that were the target of the research. In traditional alignment procedures, imperfect alignment of the spatial locations of small signals in different participants would lead to a complete loss of that signal in the average activation map. By contrast, the developed method, which was coined the 'pattern based alignment (PBA) technique' permits to find fine spatial activity patterns in averaged data. In the empirical studies presented here, the alignment problem was solved by morphing regions of interest (ROIs) to a common reference template. The PBA technique is generally applicable as long as corresponding reference points or landmarks can be defined to constrain the alignment and construct a common reference template. PBA promises to gain even more in relevance with the advent of (ultra)-high field fMRI. High-field fMRI already yields functional maps with functional patterns of uncommonly high spatial resolution, which with PBA now can be referenced and averaged across individuals without major signal loss.

The first paradigm used to investigate correlates of surface perception was dynamic brightness induction. In this paradigm, an illusory brightness effect was induced in a static grey surface in counter phase to the luminance modulations in two abutting inducer surfaces. The grey and inducer surfaces were localized by independent localizers, which produced peaks in local activity that were used as corresponding reference points for PBA of activity

patterns in V1 to V3. PBA permitted the identification of a signal in antiphase with luminance modulations in inducer surfaces, which is a likely correlate of brightness induction. This correlate varied in magnitude in accordance with the strength of brightness induction across a large set of experimental and control conditions, and was present in extrastriate cortex (stronger in V2 than in V3), but absent in V1.

Based on the findings of a correlate of brightness induction, a biologically inspired neural network of the early visual areas was built. In this network, the elementary unit was a cortical column. A matrix of columns was constructed to represent an area of the visual field, and columns were interconnected according to biologically plausible rules. By pooling network units and by applying hemodynamic constraints, activity in the model network could be converted into a modeled fMRI signal. By projecting the network generated data into the same brain space as the empirical data, a direct comparison could be made between empirical and model data using identical analysis tools. The model assumes a signal related to the physical luminance of the constant surface, which is modulated by (the dynamic changes in) local contrast information at the edges between constant grey surface and inducers. In essence, this is an interpolation model applied to brightness induction, although some of our assumptions were different from classical interpolation models. Although there is a good match between the model data and the empirical data, building additional or even very different assumptions might lead to an even better correspondence. Hence, the presently used modeling approach opens the possibility for straightforward testing of model data derived from different possible model assumptions against the empirical data. The higher the correspondence between model data and empirical data, the higher the plausibility of the mechanisms incorporated in the model. However, a relatively simple neural network model incorporating edge representations interacting with a surface spreading mechanism was already able to approach the empirical data closely.

The second paradigm used to investigate correlates of surface perception was the Troxler fading paradigm. In this paradigm, when a participant fixates a point, a peripheral figure becomes stabilized on the retina, which leads to a disappearance from awareness, and filling-in of the figure by the surrounding background. This illusion, referred to as 'perceptual filling-in' was used to probe the mechanisms underlying texture surface representation. Specifically, the question was asked whether in the cortical retinotopic representations of the interior of a physically grey figure, activity increases

## *6 Discussion*

would be found corresponding to activation by the texture. Participants in this experiment underwent fMRI scanning, and were required to fixate a centrally presented fixation point while being presented with stimuli designed to permit filling-in, and with stimuli designed to prevent the filling-in percept. To find a retinotopic correlate of the filling-in percept, a filling-in condition was compared with a filling-in prevention condition. PBA was used to aggregate activity profiles in V1 and in V2 across hemispheres and participants. We found a signal in the center of the figure's V2 representation during the filling-in condition as compared to the filling-in prevention condition. This spatially specific increase in response to the texture filling-in condition is attributed to the percept of texture in the figure region, in the absence of physical evidence for texture. This signal in V2 was closely correlated with the strength of experienced texture filling-in as measured during fMRI scanning. This signal was not present in V1.

In summary, evidence from fMRI using two different paradigms suggests that area V2 carries a signal related to the subjective perception of surfaces. The evidence provided is in good agreement with findings from recording and optical imaging studies in cats and monkeys. The fMRI studies presented here differ with previously published fMRI studies with respect to stimulus design, experimental design, and cross subject alignment. Some or all of these aspects may explain the positive findings reported in this thesis, compared to negative or opposite findings in other studies. Hence, the data provide strong support for an explicit encoding of both surface brightness and surface texture in V2. Interestingly, no such support was found in V1, suggesting a hierarchical process in which V2 computes perceptual aspects of surfaces that go beyond what is encoded in V1. The data reject 'symbolic' models that deny any explicit topographic surface encoding, and are compatible with surface interpolation models, and any other models that do explicitly represent the perceived aspect of surfaces. To more precisely test specific computational models of surface perception, further research is required.

## **Limits of the presented research**

Throughout this work, filling-in was defined as the perception of a surface feature in a region where that feature is physically absent. The empirical evidence we have obtained and the modeling we have performed show that it is plausible to assume that the neural mechanism underlying perceptual filling-in is some kind of neural interpolation process. Although less rele-

vant for brightness induction, for texture filling-in, a prominent unanswered question is which features are preserved during interpolation. The dynamic texture used was a multidimensional stimulus characterized by bars of a specific orientation and luminance, a background of lower luminance, motion, etc. Prior research (Ramachandran and Gregory, 1991) has shown that the filling-in of different features of a surface can be initiated at different moments for different surface features. In our study, we do not know which aspects of the texture are spread into the figure region during perceptual filling-in, and we do not know to what extent this spreading mechanism preserves the different aspects and statistics of the texture. Therefore, it is likely that the fMRI signal that we have called a correlate of texture filling-in is a multiplex signal consisting of temporally offset signals driven by a variety of properties of the texture.

Furthermore, we have explicitly used perceptual filling-in as a paradigm to study neuronal processes that lead to a specific visual experience of a surface. The underlying assumption is that specific sets of neurons may be involved for surface interpolation of different surface properties, through spreading in horizontally connected networks. For example, in the texture stimulus used, lateral connectivity among neurons that code brightness or color in V2 thin stripes may be a substrate for brightness or color spreading during perceptual filling-in, orientation selective neurons in V2 interstripes may contribute to spreading of information about texture element orientation, and motion selective neurons in V2 thick stripes may contribute to spreading of the dynamic aspects of the texture (Gattass et al., 1998; Sincich and Horton, 2005b, a). By contrast, Lamme, (1995); Zipser et al., (1996), have published a series of important papers (Lamme et al., 1998; Lamme and Roelfsema, 2000) showing that V1 neurons increase their activation when a large area overlaying their RFs is perceived as a figure as opposed to a ground. Their work also shows that the activation increase starts about 100ms after stimulus onset at the figure-background border, and then proceeds quickly to the middle of the figure surface (for review, Lamme and Roelfsema, 2000). Importantly, the same neurons show this effect irrespective of the manner in which the surface is defined against its background (by color, stereopsis, texture, or luminance; Lamme, 1995; Zipser et al., 1996; Levitt and Lund, 1997). Hence, the filling-in process that is revealed in this experimental paradigm may correspond to a general, contextual process that helps determine the figural status of a surface. This interpretation is consistent with the computational models that have been proposed to simulate the observed data, in which the importance of feedback

## *6 Discussion*

from high-level visual areas is emphasized (Roelfsema et al., 2002; Jehee et al., 2007). It is likely that the neural processes related to visual feature spreading across a surface, and the neural processes related to spreading of figural status across that surface interact. It will be important to start teasing apart these different forms of filling-in and their neural correlates in future studies.

## **What have we learned at the mechanistic level?**

Although the presented empirical work has led to strong support for explicit encoding of the visually perceived aspects of surfaces across their spatial extent, the fMRI data we have collected say little about the neuronal circuitry that produce the fMRI responses. Based on neurophysiological and optical imaging studies, it is possible to formulate specific hypotheses as to the contributions of different subsystems within V1 and V2. Evidence reviewed in the Introduction suggests that a pathway involving V1 blobs and V2 thin stripes may play a privileged role in the spreading of visual surface features, while a pathway involving V1 interblobs and V2 inter and thick stripes may play a privileged role in the construction of boundary representations (Sincich and Horton, 2005b, a). In addition, anatomical evidence shows that lateral connectivity that could contribute to spreading is especially prominent in superficial layers of cortex (Gilbert, 1977; Gilbert and Wiesel, 1979, 1985, 1989; Gilbert et al., 1990; Gilbert and Wiesel, 1990, 1992; Grinvald et al., 1994; Tanifuji et al., 1994), and that the results of global, contextual operations in high-level areas may re-enter preferably through deep layers of V1 and V2 (Angelucci and Bressloff, 2006). The above mentioned hypothetical assignment of different functions to different compartments in V1 and V2 leads to specific predictions for filling-in in a Troxler fading paradigm. For example, during stabilization there may be a decline of activity in the V1 interblob – V2 thick/inter stripe subsystem. Furthermore, only during time periods in which this decline is sufficient (during accurate fixation), one may observe a lateral spreading process involving the V2 thin stripe system. In general, measuring activity within the different subsystems in V1 and V2, and measuring functional connectivity (Goebel et al., 2003; Roebroeck et al., 2005, 2009) across these subsystems and across cortical layers will be a prerequisite to test specific aspects of computational models, such as the interactions between boundary and surface processing systems hypothesized in interpolation theory.

## **Future perspectives**

To successfully test detailed mechanistic hypotheses about differential contributions of functional subcompartments in early visual areas, there are two basic requirements. First, it would be beneficial to maximize the strength of the illusion used to probe surface perception. This may be achieved by improving stimulus design. Specifically, in the case of filling-in in the Troxler paradigm, we faced the difficulty to obtain a large signal related to the illusion, because of the contradictory requirements to maximize figure size to compensate for limits in fMRI spatial resolution, and to minimize figure size to promote the illusion of filling-in. This problem might be addressed by increasing the strength of the illusion by artificially increasing the stabilization of the image on the retina. In principle, this can be achieved by using fixation errors for online correction of the figure's position. This requires a stimulus presentation system with a high refresh rate combined with a high spatial and temporal resolution eye tracking system. If this form of artificial image stabilization can be achieved, it may become possible to fill in figures of significant size in (para)foveal vision, which is expected to lead to neural interpolation across large regions of cortex because of cortical magnification. Although the implementation of artificial image stabilization in an MR environment involves significant hardware challenges, the potential experimental benefits warrant exploring that approach.

Second, it would be beneficial to enhance the spatial and temporal resolution at which functional data are collected. Remarkably, although addressing detailed spatial and temporal aspects of neural signals related to surface perception is within the possibilities of currently available neurophysiological techniques in animals (e.g., Schroeder et al., 1998; Yoon et al., 2000; Rubehn et al., 2009), this possibility so far has been taken little advantage of in studies of surface perception. However, with the advent of (ultra) high-field imaging, and optimization of imaging at 3T, functional neuroimaging suited for resolving anatomical layers, functional subdomains within cortical areas, and columns may become feasible (Yacoub et al., 2008; Bissig and Berkowitz, 2009; Koopmans et al., 2010). Thanks to these developments in the field of fMRI, detailed experiments testing mechanistic hypotheses about surface perception may soon be within reach in human experimental subjects.

## 6 Discussion

## References

- Angelucci, A., and Bressloff, P.C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Prog Brain Res* 154, 93-120.
- Bissig, D., and Berkowitz, B.A. (2009). Manganese-enhanced MRI of layer-specific activity in the visual cortex from awake and free-moving rats. *Neuroimage* 44, 627-635.
- Gattass, R., Pessoa, L.A., De Weerd, P., and Fiorani, M. (1998). Filling-in in topographically organized distributed networks. *Proceedings of the Brazilian Academy of Sciences* 70, 1-11.
- Gilbert, C.D. (1977). Laminar differences in receptive field properties of cells in cat primary visual cortex. *J Physiol* 268, 391-421.
- Gilbert, C.D., Hirsch, J.A., and Wiesel, T.N. (1990). Lateral interactions in visual cortex. *Cold Spring Harb Symp Quant Biol* 55, 663-677.
- Gilbert, C.D., and Wiesel, T.N. (1979). Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. *Nature* 280, 120-125.
- Gilbert, C.D., and Wiesel, T.N. (1985). Intrinsic connectivity and receptive field properties in visual cortex. *Vision Res* 25, 365-374.
- Gilbert, C.D., and Wiesel, T.N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J Neurosci* 9, 2432-2442.
- Gilbert, C.D., and Wiesel, T.N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Res* 30, 1689-1701.
- Gilbert, C.D., and Wiesel, T.N. (1992). Receptive field dynamics in adult primary visual cortex. *Nature* 356, 150-152.
- Goebel, R., Roebroeck, A., Kim, D.S., and Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn Reson Imaging* 21, 1251-1261.
- Grinvald, A., Lieke, E.E., Frostig, R.D., and Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *J Neurosci* 14, 2545-2568.
- Jehee, J.F., Lamme, V.A., and Roelfsema, P.R. (2007). Boundary assignment in a recurrent network architecture. *Vision Res* 47, 1153-1165.
- Koopmans, P.J., Barth, M., and Norris, D.G. (2010). Layer-specific BOLD activation in human V1. *Hum Brain Mapp* 31, 1297-1304.
- Lamme, V.A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15, 1605-1615.
- Lamme, V.A., and Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci* 23, 571-579.
- Lamme, V.A., Super, H., and Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Curr Opin Neurobiol* 8, 529-535.

## References

- Levitt, J.B., and Lund, J.S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature* 387, 73-76.
- Ramachandran, V.S., and Gregory, R.L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature* 350, 699-702.
- Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage* 25, 230-242.
- Roebroeck, A., Formisano, E., and Goebel, R. (2009). The identification of interacting networks in the brain using fMRI: Model selection, causality and deconvolution. *Neuroimage*.
- Roelfsema, P.R., Lamme, V.A., Spekreijse, H., and Bosch, H. (2002). Figure-ground segregation in a recurrent network architecture. *J Cogn Neurosci* 14, 525-537.
- Rubehn, B., Bosman, C., Oostenveld, R., Fries, P., and Stieglitz, T. (2009). A MEMS-based flexible multichannel ECoG-electrode array. *J Neural Eng* 6, 036003.
- Schroeder, C.E., Mehta, A.D., and Givre, S.J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cereb Cortex* 8, 575-592.
- Sincich, L.C., and Horton, J.C. (2005a). The circuitry of V1 and V2: integration of color, form, and motion. *Annu Rev Neurosci* 28, 303-326.
- Sincich, L.C., and Horton, J.C. (2005b). Input to V2 thin stripes arises from V1 cytochrome oxidase patches. *J Neurosci* 25, 10087-10093.
- Tanifuji, M., Sugiyama, T., and Murase, K. (1994). Horizontal propagation of excitation in rat visual cortical slices revealed by optical imaging. *Science* 266, 1057-1059.
- Yacoub, E., Harel, N., and Ugurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proc Natl Acad Sci U S A* 105, 10607-10612.
- Yoon, T.H., Hwang, E.J., Shin, D.Y., Park, S.I., Oh, S.J., Jung, S.C., Shin, H.C., and Kim, S.J. (2000). A micromachined silicon depth probe for multichannel neural recording. *IEEE Trans Biomed Eng* 47, 1082-1087.
- Zipser, K., Lamme, V.A., and Schiller, P.H. (1996). Contextual modulation in primary visual cortex. *J Neurosci* 16, 7376-7389.

## **7 Samenvatting**

## 7 Samenvatting

Een groot deel van de energie beschikbaar voor visuele verwerking in het brein wordt gewijd aan de reconstructie van de grenzen tussen vlakken gebaseerd op fragmentarische invoer afkomstig uit de retina. De processen onderliggend aan de reconstructie van deze grenzen zijn uitvoerig bestudeerd, er heert een consensus dat neurale interpolatie processen bijdragen aan het reconstructie proces. Desalniettemin is het signaal afkomstig vanuit een homogeen vlak klein, en stellen computationele modellen en empirische studies in vraag wat de relevantie van dit kleine signaal met betrekking tot de subjectieve perceptie van vlakken is. De in deze thesis beschreven studies hebben als doel inzicht te scheppen in het debat over het al dan niet bestaan van een signaal in de retinotopische representatie van een vlak dat gelieerd is aan de perceptie van dat vlak. We hebben deze vraag onderzocht in twee psychofysische paradigma's waarin de perceptie van vlakken werd gemanipuleerd in afwezigheid van fysische veranderingen. We gebruiken fMRI beeldvorming om in deze paradigma's activiteit te meten in de visuele cortex.

Een noodzaak om een antwoord op de hoofdvraagstelling te kunnen formuleren was de ontwikkeling van een methode voor het efficient uitlijnen van functionele beeldvormingsdata. Het doel van deze methode is het signaalverlies ten gevolge van individuele verschillen te minimaliseren wanneer proefpersoongegevens gecombineerd worden door middel van middeling. Deze methodologische ontwikkeling is hoogstnodig in het licht van de hypothese dat de magnitude van het signaal afkomstig van vlakken klein is. Het gebruik van de traditionele uitlijn methoden resulteert veelal in de imperfecte uitlijning van spatiële locaties, afkomstig van meerdere proefpersonen, waaruit de kleine signalen ontstaan en leidt tot algeheel signaalverlies in gemiddelde activatieoverzichten. In sterk contrast hiertoe staat dat de ontwikkelde patroon gebaseerde uitlijn methode (PGU) ertoe in staat is om fijnschalige activiteitspatronen te ontdekken in gemiddelde proefpersoongegevens. In de empirische studies hier gepresenteerd wordt het uitlijningsprobleem opgelost door het vervormen van de bestudeerde regios in het brein naar een gemeenschappelijk referentie sjabloon. De PGU techniek is algemeen toepasbaar wanneer de mogelijkheid bestaat om corresponderende referentiepunten of andere oriëntatiepunten tussen hersenenregios in proefpersonen te definiëren om het uitlijnprocess te begeleiden en om een gemeenschappelijk referentie sjabloon te vormen. PGU belooft zelfs meer relevant te worden door de opkomst van ultra hoge veld fMRI. Ultra hoge veld fMRI is reeds in staat om functionele activiteitsoverzichten op te leveren van een ongemeen hoge spatiële resolutie die door middel van PGU uitgelijnd kan worden waardoor data overeen subjecten gemiddeld kan worden zonder noemenswaardig signaalverlies.

Het eerste paradigma dat gebruikt werd om de correlaten van vlak perceptie te bestuderen was een dynamische helderheids inductie paradigma. Met dit paradigma kon een illusoir helderheids effect geïnduceerd worden in een statisch grijs vlak dat in tegenfase is met de luminantiemodulaties in twee aangrenzende inducerende vlakken. Het grijze vlak en de inducerende vlakken werden retinotopisch gelocaliseerd door onafhankelijke localisatie stimuli. De hieruit resulterende lokale activatiepieken werden gebruikt als corresponderende referentiepunten ten behoeve van de PGU van activiteitspatronen in V1 tot V3. Door middel van PGU werd een signaal geïdentificeerd dat zich in tegenfase verhield in vergelijking met het signaal ten gevolge van de luminantiemodulaties in de inducerende vlakken. Het tegenfase signaal is hoogstwaarschijnlijk een correlaat van helderheidsinductie want de magnitude van het signaal covarieerde met de door proefpersonen ingeschatte sterkte van het illusoire helderheids effect doorheen een reeks van experimentele en controle experimenten. Het illusoire helderheids effect was meest prominent aanwezig in V2, zwakker in V3 maar afwezig in V1.

Gebaseerd op de observatie van een correlaat van helderheidsinductie werd een biologisch geïnspireerd neurale netwerk gebouwd. In dit netwerk werden de elementaire eenheden aan corticale kolommen gelijkgesteld. Een matrix van kolommen werd zo geconstrueerd dat deze een deel van het visuele veld vertegenwoordigde, deze kolommen werden onderling verbonden aan de hand van biologisch plausibele regels. Door het samenvoegen van netwerkeenheden en door het toepassen van een hemodynamische filter op de signaal van de samengevoegde netwerkeenheden kon het netwerksignaal naar een fMRI signaal geconverteerd worden. Door het projecteren van de door het netwerk gegenereerde data naar hetzelfde breingebaseerd gemeenschappelijk referentiekader als de empirische

data kon op beide datasets een gelijke analysemethode gebruikt worden. Het netwerk model veronderstelt dat een signaal gerelateerd aan de luminantie van het statische vlak gemoduleerd wordt door de dynamische veranderingen van het locale contrast aan de grenzen tussen het statische vlak en de inducerende vlakken. Het neurale netwerk is gebouwd op principes ingegeven door de interpolatietheorie van vlakken, hoewel sommige van onze assumpties verschillen van de inzichten vervat in klassieke interpolatie modellen. Ofschoon een goede overeenkomst tussen model data en empirische data werd vastgesteld is het mogelijk om additionele of compleet nieuwe voorwaarden in het model te bouwen die tot een verbeterde overeenkomst kunnen leiden. Dus de modelleeraanpak opent de mogelijkheid om op een simpele manier directe vergelijkingen te maken tussen empirische data en modeldata gebaseerd op verschillende assumpties. Hoe hoger de overeenkomst tussen modeldata en empirische data, hoe hoger de plausibiliteit van het mechanisme vervat in het model. Desalniettemin kan modeldata afkomstig van een relatief simpel neurale netwerk model met daarin vervat grensrepresentaties die interacteren met vlakspreidingsmechanismen de empirische data dichtbij benaderen.

Het tweede paradigma dat gebruikt werd om de correlaten van vlak perceptie te bestuderen was een Troxlervervagings paradigma. In dit paradigma wordt een stimulus gebruikt die is opgebouwd uit een figuur dat omringd is door dynamische textuur. Doordat een persoon zijn blik fixeert op een punt, wordt een perifere figuur gestabiliseerd op de retina wat ertoe leidt dat de figuur uit het bewustzijn van de persoon verdwijnt en dat de figuur ingevuld wordt door de dynamische textuur. Deze illusie die perceptuele invulling genoemd wordt werd gebruikt om de mechanismen onderliggend aan de representatie van textuurvlakken ten gronde te onderzoeken. Meer bepaald was de vraagstelling of de activiteit tijdens invulling van een grijze figuur zou overeenkomen met de activatie door textuur in de retinotopische representatie van deze figuur. Deelnemers in dit experiment ondergingen fMRI beeldvorming. Hun instructie was om hun blik op een centraal gepresenteerd fixatiepunt te fixeren terwijl een stimulus die perceptuele invulling toelaat aan hen werd gepresenteerd. In een controle-experiment werd hen gevraagd te fixeren terwijl een stimulus die perceptuele invulling tegenging gepresenteerd werd. Om een retinotopisch correlaat van het perceptuele invullings percept vast te stellen werd een vergelijking gemaakt van de activiteit in de figuur regio in het brein tussen de conditie die filling-in toeliet en de conditie die filling-in tegenging. PGU werd toegepast op deze data zodat activiteits profielen van V1 en V2 geaggregateerd konden worden over hersenhemisferen en proefpersonen. Door de vergelijking van de signalen in de conditie die filling-in faciliteerde en de conditie die filling-in tegenging vonden wij een signaal in het centrum van de cortical figuurrepresentatie in V2 dat aan perceptuele invulling geattributeerd kan worden. Van deze plaatsspecifieke activiteitsstijging in reactie op de textuur filling-in conditie wordt aangenomen dat deze een gevolg is van het textuurpercept in de figuur regio zonder dat er daar fysieke textuur aanwezig is. Het signaal in V2 was sterk gecorreleerd met de sterke van het illusoire percept als aangegeven door de proefpersonen tijdens de fMRI beeldvorming. Dit signaal was niet aanwezig in V1.

Samenvattend wordt gesuggereerd dat ingegeven door bewijslast afkomstig van fMRI in twee verschillende paradigmas, een signaal afkomstig uit V2 gerelateerd is aan de subjectieve waarneming van vlakken. Het hier verzamelde bewijsmateriaal komt sterk overeen met bevindingen afkomstig van cel afleidingen en optische beeldvormingsexperimenten uitgevoerd in katten en primaten. De gepresenteerde fMRI experimenten verschillen van eerder gepubliceerde fMRI studies wat betreft de gebruikte stimuli, het experimentele ontwerp en door de PGU techniek. Enkele, of al deze verschilpunten kunnen de positieve bevindingen die gerapporteerd worden in deze thesis verklaren wanneer een vergelijking gemaakt wordt met de tegengestelde en nul-bevindingen in andere experimenten. Dus deze leveren sterke ondersteuning aan de theorie dat zowel vlakhelderheid en vlaktextuur explicet geïncodeerd worden in V2. Interessant is dat deze bevinding niet opgaat voor V1, dit suggereert een hiërarchische verwerkingsproces waarin V2 verantwoordelijk is voor computaties met betrekking tot de perceptuele aspecten van vlakken die de rol van V1 overstijgen. Op grond van deze data kunnen 'symbolische' modellen die een expliciete topografische codering verwerpen afgewezen worden, de gevonden data zijn echter wel compatibel met vlak interpolatie modellen en eender welke klasse van modellen die poneert dat het perceptuele aspect van vlakken explicet gecodeerd wordt. Om te

## *7 Samenvatting*

discrimineren tussen de resterende computationele modellen is verder onderzoek noodzakelijk.

# **Curriculum Vitae**

## *7 Samenvatting*

Bert Jans was born on December 6th in Tongeren, Belgium. In 2000 he completed his secondary school education at the Technisch Instituut Onze Lieve Vrouw, in Tongeren. In 2001 He started the Bachelor Psychology at the University of Maastricht. He obtained his master degree in Psychology in August 2005. He started a Ph.D project in October 2005 under the supervision of Prof. Dr. Peter De Weerd and Prof. Dr. Rainer Goebel at the department of Cognitive Neuroscience in the Faculty of Psychology and Neuroscience at Maastricht University.