Processing of second-order, contrast-modulated stimuli in mouse visual cortex

Dissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
der Mathematisch-Naturwissenschaftlichen Fakultät
und
der Medizinischen Fakultät
der Eberhard-Karls-Universität Tübingen

vorgelegt
von

Zeinab Khashkhodaei
aus Shiraz, Iran

März, 2017
Tag der mündlichen Prüfung: 16.03.2017

Dekan der Math.-Nat. Fakultät: Prof. Dr. W. Rosenstiel
Dekan der Medizinischen Fakultät: Prof. Dr. I. B. Autenrieth

1. Berichterstatter: Prof. Dr. / PD Dr. / Dr. Laura Busse
2. Berichterstatter: Prof. Dr. / PD Dr. / Dr. Thomas Euler

Prüfungskommission: Prof. Dr. / PD Dr. / Dr. Laura Busse

Prof. Dr. / PD Dr. / Dr. Thomas Euler
Prof. Dr. / PD Dr. / Dr. Uwe Ilg
Prof. Dr. / PD Dr. / Dr. Cornelius Schwarz
Erklärung / Declaration:

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:

Processing of second-order, contrast-modulated stimuli in mouse visual cortex

selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

I hereby declare that I have produced the work entitled “……”, submitted for the award of a doctorate, on my own (without external help), have used only the sources and aids indicated and have marked passages included from other works, whether verbatim or in content, as such. I swear upon oath that these statements are true and that I have not concealed anything. I am aware that making a false declaration under oath is punishable by a term of imprisonment of up to three years or by a fine.

Tübingen, den 26.06.2017

Zeinab Khaustkhodaei

Datum / Date

Unterschrift /Signature
Table of Contents

Summary.......................................................................................................................... 4

Introduction ....................................................................................................................... 6
  Central visual pathways................................................................................................. 6
  Hierarchy of visual processing ...................................................................................... 6
  Gradually increasing invariance .................................................................................. 7
  Form-cue invariance .................................................................................................... 8
    First-order and second-order stimuli ......................................................................... 8
  Neuronal responses to second-order stimuli.................................................................. 9
  Cortical organization of responsive neurons to second-order stimuli......................... 11
  Potential model for second-order processing .............................................................. 12
    Filter-Rectify-Filter model ....................................................................................... 12
    Surround suppression mechanism ......................................................................... 13
  Mice as a popular model system for visual processing ............................................... 14
    Functional organization of mouse higher visual areas .............................................. 14
    Can mice vision rely on more than luminance cues? ............................................. 15
    Area LM might be a suitable area to investigate second-order processing ............. 15
  Can mice use second-order stimuli to guide visual perception in a cue-invariant way? 16

Material & Methods ....................................................................................................... 18
  Luminance-modulated and contrast-modulated gratings ............................................. 18
  Analysis of visual stimuli ............................................................................................. 18
  Surgical preparation for head-fixed visual behavior .................................................... 19
  Orientation discrimination task and analysis of behavioral data ................................ 19
  Surgical preparation for electrophysiological recordings during wakefulness .......... 20
  Surgical preparation for electrophysiological recordings during anesthesia .......... 20
  Visual stimuli for electrophysiological recordings ..................................................... 21
  Analysis of electrophysiological data ......................................................................... 21
    Unit extraction and spike sorting ........................................................................... 21
    Analysis of tuning ...................................................................................................... 21
  Comparison of responses to luminance-modulated and contrast-modulated gratings .. 22

Histology .......................................................................................................................... 22

Results ............................................................................................................................... 24
  Stimuli............................................................................................................................. 24
  Orientation discrimination learning for luminance-modulated gratings .................... 25
  Cue-invariant generalization of orientation discrimination .......................................... 26
    Identification of areas V1 and LM based on mirrored retinotopic representation of 26
      azimuth ..................................................................................................................... 28
  Responses to contrast-modulated gratings are weaker and less selective ................. 29
    Responses to low-frequency noise contrast-modulated gratings ............................. 30
    Responses to RMS matched contrast luminance-modulated gratings .................... 31
    Responses to high-frequency noise contrast-modulated gratings .......................... 33
    Preferred orientations for luminance-modulated and contrast-modulated gratings are 34
      broadly similar ........................................................................................................ 34
    Responses to contrast-modulated gratings during anesthesia ................................ 35

Discussion ......................................................................................................................... 39
  Choice of paradigm ...................................................................................................... 39
  Why did I choose classical conditioning? .................................................................. 39
  Can classical conditioning paradigm reveal the limits of visual performance? .......... 40
Underlying mechanism of second-order.................................................................41
Potential role of LM in perception of texture boundaries .........................................41
Are the other mouse extrastriate areas more strongly responsive to second-order
stimuli?.........................................................................................................................42
It is also currently unknown whether responses to second-order gratings are stronger
during task performance. ..........................................................................................43
Conclusion..................................................................................................................44
Acknowledgment.......................................................................................................45
References..................................................................................................................47
FIGURE 1. TWO PARALLEL PATHWAYS OF VISUAL PROCESSING IN PRIMATES ................................................................. 6
FIGURE 2. IN THE HIERARCHY OF VENTRAL STREAM (E.G. IN V4 AND IT AREAS), NEURAL SELECTIVITY IS TOLERANT TO CHANGES IN LOW-LEVEL VISUAL FEATURES ................................................................. 8
FIGURE 3. EXAMPLES OF FIRST-ORDER AND SECOND-ORDER STIMULI IN THE NATURE ............................................... 8
FIGURE 4. EXAMPLES OF SECOND-ORDER STIMULI USED IN PREVIOUS STUDIES .......................................................... 9
FIGURE 5. THE PREFERRED SPATIAL FREQUENCIES OF CARRIER (RED) ARE AROUND 8 TIMES THAN THOSE FOR ENVELOPE (BLUE) ........................................................................................................... 10
FIGURE 6. THERE IS NO FIXED RELATIONSHIP BETWEEN CARRIER (RED) AND ENVELOPE (BLUE) ORIENTATIONS ............................................................................................................................... 11
FIGURE 7. LUMINANCE GRATINGS PRESENTED IN AN OPTIMAL SPATIAL FREQUENCY AND ORIENTATION EVOKE MAXIMAL NEURAL RESPONSES .................................................................................. 12
FIGURE 8. LINEAR MECHANISM CANNOT EXPLAIN THE SECOND-ORDER RESPONSES, UNLIKE TO FIRST-ORDER ........................................................................................................................................... 12
FIGURE 9. TWO-STAGE MODEL SUGGESTED FOR SECOND-ORDER PROCESSING .................................................................. 13
FIGURE 10. EXTRASTRIATE VISUAL AREAS IN THE PRIMATE AND MOUSE ........................................................................... 15
FIGURE 11. BEHAVIOR SETUP, RECORDINGS, AND ISOLATED SINGLE NEURONS ........................................................................ 23
FIGURE 12. ANALYSIS OF VISUAL STIMULI ............................................................................................................................. 25
FIGURE 13. ORIENTATION DISCRIMINATION FOR LGs, ..................................................................................................................... 26
FIGURE 14. BEHAVIORAL PERFORMANCE FOR CGs AND CONTROL CONDITIONS ....................................................................... 28
FIGURE 15. IDENTIFICATION OF VISUAL AREAS V1 AND LM BY MIRRORED RETINOTOPIC PROGRESSION OF AZIMUTH
FIGURE 5. EXAMPLE NEURONS UNRESPONSIVE TO CGs DESPITE SIGNIFICANT RESPONSES TO LGs .......................... 30
FIGURE 6. DISTRIBUTION OF RESPONSIVE AND UNRESPONSIVE NEURONS TO CGs WITH LOW-FREQUENCY CARRIER IN AREA V1 AND LM ..................................................................................................................... 31
FIGURE 7. RESPONSES TO CGs WITH LOW-FREQUENCY CARRIER IN MOUSE VISUAL CORTEX .............................................. 31
FIGURE 8. DISTRIBUTION OF RESPONSIVE AND UNRESPONSIVE NEURONS TO CGs WITH LOW-FREQUENCY CARRIER IN RMS-MATCHED CONTRAST LGs EXPERIMENTS, IN AREA V1 AND LM ........................................................................................................................................ 32
FIGURE 9. COMPARISON OF RESPONSES TO LGs MATCHED IN RMS CONTRAST AND CGs ................................................................. 33
FIGURE 10. DISTRIBUTION OF RESPONSIVE AND UNRESPONSIVE NEURONS TO CGs WITH HIGH-FREQUENCY CARRIER IN AREA V1 AND LM ................................................................................................................. 34
FIGURE 11. RESPONSES TO CGs WITH HIGH-FREQUENCY CARRIER ............................................................................................... 34
FIGURE 12. COMPARISON OF PREFERRED ORIENTATION FOR CONTRAST-MODULATED VERSUS LGs .......................... 35
FIGURE 13. EXAMPLE NEURONS UNRESPONSIVE TO CGs DESPITE SIGNIFICANT RESPONSES TO LGs DURING ANESTHESIA ...................................................................................................................... 35
FIGURE 14. DISTRIBUTION OF RESPONSIVE AND UNRESPONSIVE NEURONS TO CGs WITH LOW-FREQUENCY CARRIER IN AREA V1 AND LM DURING ANESTHESIA .................................................................................. 36
FIGURE 15. EXAMPLE NEURONS UNRESPONSIVE TO CGs DESPITE SIGNIFICANT RESPONSES TO LGs DURING ANESTHESIA ...................................................................................................................... 37
Summary

Visual processing along the primate ventral stream takes place across a hierarchy of areas, characterized by an increase in both complexity of neuronal preferences and invariance to changes of low-level stimulus attributes. A basic type of invariance is form-cue invariance, where neurons have similar preferences in response to first-order stimuli, defined by changes in luminance, and global features of second-order stimuli, defined by changes in texture or contrast. Whether in mice, a now popular model system for early visual processing, visual perception can be guided by second-order stimuli is currently unknown.

In this project, we asked whether mice can use second-order stimuli to guide visual perception in a cue-invariant way and assessed potential cue-invariant representations of stimulus orientation in two areas of mouse visual cortex. For both behavioral and electrophysiological experiments, we used a common set of luminance-modulated gratings (LGs) and contrast-modulated gratings (CGs), obtained by multiplying a contrast envelope with a noise carrier. We created two types of CG stimuli that differed in the Fourier energy distribution of the carrier: high-frequency noise and low-frequency noise. To examine a potential effect of different contrast of LGs and CGs, we matched the root-mean-square (RMS) contrast between first- and second-order gratings, in which case c was 0.335 for the LGs.

We tested whether mice can generalize orientation discrimination learned with first-order, LGs to various untrained second-order, CGs. We first trained head-fixed mice in a classical conditioning paradigm to perform a coarse orientation discrimination on LGs. Once the animal reached stable and reliable orientation discrimination performance, we replaced LGs with CGs with low-frequency noise carriers to test the generalization of orientation discrimination to second-order gratings. We found that mice, after learning a coarse orientation discrimination involving only LGs, could readily generalize orientation discrimination to CGs, albeit with a substantial drop in performance. Then, we wondered whether the overall lower performance for CGs was related to their lower RMS contrast compared with LGs. To test this hypothesis, we probed mice with LGs whose RMS contrast was lowered to match that of CGs. We found that mice could perform well during the orientation discrimination task for LGs matched in RMS contrast. Indeed, across all animals tested, performance was similar for both levels of contrast. Finally, we tested mice with CGs with high-frequency noise carriers. Again, mice could see this type of CG and importantly, they could also discriminate between the two grating orientations, albeit performance was again considerably lower compared with that for LGs. Together, these results demonstrate that mice can use second-order stimuli to guide visual perception.

We performed extracellular recordings, both during anesthesia and wakefulness, in mouse areas V1 and LM, where we compared orientation tuning curves to LGs and CGs. We found that neurons in area V1 and LM were less responsive and less selective to CGs than to LGs, both during anesthesia and wakefulness. Interestingly, this reduction was particularly prominent during anesthesia. We wondered whether our finding of weaker responses and broader orientation tuning for CG than LG responses could be explained by the lower RMS contrast of CGs. Thus, we performed control experiments, in which we measured responses to LGs that were matched in RMS contrast to the CGs. We found that neurons in both visual areas still responded more weakly to CGs than LGs but this reduction in responsiveness was less pronounced compared with conditions with full
contrast LGs. Similar to our results with full-contrast LGs, orientation selectivity also decreased considerably between LGs matched in RMS contrast and CGs. Indeed, orientation selectivity did not differ significantly between responses to full-contrast and reduced-contrast LGs. Together, the reduced RMS contrast of CGs might contribute to the reduction in peak responsiveness to CGs but cannot account for the poorer orientation selectivity for CGs. We also investigated the underlying mechanism of responses to second-order stimuli in awake recordings: we tested orientation tunings of both V1 and area LM in response to a high-frequency noise CGs, in which the noise carrier’s spatial frequency distribution was concentrated beyond the passband of many V1 and LM neurons (Marshel et al., 2011). We first observed that less than half of the recorded neurons with significant responses to LGs also had visually evoked activity to CGs with high-frequency noise. This fraction of responsive neurons was considerably lower compared to that obtained for CGs with low-frequency noise. Interestingly, the difference in responsiveness between the two types of CG stimuli was stronger for area V1 than LM. To examine whether the CG representation might contribute toward cue-invariant perception of stimulus orientation, we also compared the neurons’ preferred orientation, separately for each grating type. We found that preferred orientations for CGs with low-frequency noise and LGs were correlated for both areas V1 and LM. Interestingly, for CGs with high-frequency noise, the distribution of differences in preferred orientation was non-uniform only for area LM and preferred orientations were only correlated for area LM. Together, the broad similarity of preferred orientations between grating types provides some evidence for a coarse cue-invariance, which might in turn be part of the neural basis for perceptual generalization of orientation discrimination.
Introduction

Central visual pathways
The last 50 years have seen major advances in our understanding of the visual system, and in particular of the visual cortex. Responses of neurons in the mammalian visual cortex were first thoroughly quantified by Hubel and Wiesel, more than 50 years ago (Hubel and Wiesel, 1959). Today, we know that the processing of visual information is performed to a large extent in the visual cortex, which is subdivided into distinct areas: in the primate, there are approximately 32 separate neocortical areas involved in visual processing (Felleman and Van Essen, 1991). These areas are organized into two broadly segregated processing streams: the ventral stream for object recognition and the dorsal stream for object localization. Each stream consists of multiple areas beyond V1 and has V1 as a primary input source (Mishkin and Ungerleider, 1982; Ungerleider and Haxby, 1994) (Figure 1).

Hierarchy of visual processing
In primates, visual processing along the ventral stream takes place across a hierarchy of areas. In this hierarchy the complexity of neural preferences increases, such that V1 neurons prefer orientated edges (Hubel and Wiesel, 1959, 1962; De Valois et al., 1982), V2 neurons respond to contours, textures and combination of orientations (Anzai et al., 2007; Willmore et al., 2010; Freeman et al., 2013), V4 neurons process curvatures (Pasupathy and Connor, 1999) and finally neurons in IT, the final stage of the ventral
stream, prefer complex objects such as faces (Desimone et al., 1984; Logothetis and Sheinberg, 1996; Tanaka, 1996).

The selectivity of V1 neurons was reported, for the first time, by Hubel and Wiesel (Hubel and Wiesel, 1959). In the seminal experiments of Hubel and Wiesel, they discovered that cells in cat primary visual cortex tended to respond in one of two ways. The first group, which they termed “simple cells”, responds more when the light is presented in a certain direction and these responses can be predicted based on the excitatory and inhibitory organization of their receptive fields. However, the second group, which they named “complex cells”, does not have fixed inhibitory and excitatory subregions in their receptive fields and the direction selectivity of these neurons cannot be predicted based on their receptive fields organizations. One step further in the hierarchy, V2 neurons respond to combinations of similar and dissimilar orientations.

The property of having different orientation preferences in receptive field subregions gives V2 neurons the capability to process curves and angles (Anzai et al., 2007), despite the fact V2 neurons do not show more than one peak in their orientation tuning curve when probed with bars or gratings. Further evidence for gradually increasing complexity comes from a study comparing V1 and V2 neuronal responses to natural images. Whereas most of V2 neurons respond to natural images stronger than to control stimuli, this modulation effect is minimal in V1 neurons (Freeman et al., 2013).

V4 neurons process even more complicated visual features and prepare the information for shape recognition in the next stages. The modulation depth of V4 neurons by contour features is much stronger than by edges and bars. Moreover, a subpopulation of V4 neurons are selective for contour features and show a bias toward convex contours (Pasupathy and Connor, 1999). Finally, in IT, the final stage of ventral stream, neurons respond more to complex than simple images such that a subpopulation of IT neurons is specialized to recognize faces (Desimone et al., 1984; Logothetis and Sheinberg, 1996; Tanaka, 1996).

**Gradually increasing invariance**

Another example of increasing the complexity of visual information processing along ventral stream is the tolerance to changes in low-level stimulus features in object recognition. Previous studies have shown that IT neurons conserve their selectivity despite changes in position, scale and context of an image. As depicted in Figure 2, monkeys were shown different images, where each individual image varies in scale, position and context, and the ability of V4 and IT neurons to generalize their selectivity was tested. Many neurons in V4 and IT maintain their shape selectivity across changes in position, size and context (Schwartz et al., 1983; Rust and DiCarlo, 2010). Another study demonstrated that area IT, which has overlapping columns forming a continuous mapping of complex features, is able to produce the image of an object, invariant to different viewing angles and illumination conditions (Tanaka et al., 1996). Moreover, the tolerance increases as visual information travels from V4 to IT along ventral stream, and therefore neurons become more tolerant to changes in low-level stimulus attributes (Rust and DiCarlo, 2010) (Figure 2).
Form-cue invariance

One basic type of invariance is form-cue invariance (Baker and Mareschal, 2001), where neurons have similar preferences whether driven by first-order or second-order stimuli.

First-order and second-order stimuli

In order to perform object recognition, our visual system needs to first analyze the local features of natural images, which are due to surface properties, illumination, and boundaries. It has been established that neurons in early visual system encode the local changes in luminance. However, there are other types of objects, which are segregated from their background with other cues than luminance and our visual system is still able to perceive them. In the natural image in Figure 3, we recognize the tree from the grassy background and the grass from its reflection in the water with two different mechanisms: in the first case the luminance is the local attribute change between tree and grassy background, whereas the grass is distinguished from its reflection by a difference in contrast. Those stimuli whose principal features are characterized by changes in luminance are referred to as first-order stimuli; sine wave gratings are a
common type of first-order stimuli and have been used in many studies. Visual stimuli, which are characterized by variations in other visual cues than luminance, such as contrast, texture and color are defined as second-order stimuli. By now, different types of second-order stimuli have been used to investigate second-order processing. **Figure 4** illustrates a few types of second-order stimuli. One of the most common types of second-order stimuli is a contrast-modulated grating (CG). CGs consist of two elements: a noise texture, called a carrier, and a low spatial frequency grating, called an envelope. In CGs the contrast of carrier changes by envelope.

**Neuronal responses to second-order stimuli**

Neural responses to second-order motion was reported first by Albright 1992, who showed that neurons in area MT/V5 have the same preferred direction to LGs and moving dynamic noise on a static noise background (Form-cue invariance) (Albright, 1992). So far, cue-invariant neurons have been observed in primate V2 (Li et al., 2014) (but see El-Shamayleh and Movshon, 2011; An et al., 2014), MT (Albright, 1992) and IT (Sary et al., 1993), and cat area 18 (Zhou and Baker, 1994; Leventhal et al., 1998; Mareschal and Baker, 1998a, b; Zhan and Baker, 2006; Song and Baker, 2007).

In most of these electrophysiological and anatomical studies, the second-order responsive neurons showed weaker responses to second-order than to first-order stimuli (Mareschal and Baker, 1998a; Zhan and Baker, 2006; Li et al., 2014). For instance, optical imaging in area 18 of cats has shown that the population responses to both contrast-contour and illusory-contour stimuli, which are two types of second-order stimuli, were approximately 20% of those to LGs (Zhan and Baker, 2006). Similar results have been obtained in electrophysiological studies conducted in monkeys and cats, such that only one-third of V2 neurons responded to CGs and these responses were weaker than those to first-order stimuli (Mareschal and Baker, 1998a; Li and Baker, 2012; Li et al., 2014). Therefore, cue-invariant neurons have been shown to be typically less responsive to second-order than to first-order stimuli.

The properties of both global and local features of second-order stimuli have been studied in details. Spatial frequency and orientation of both the carrier and the envelope have been shown to determine second-order responses. Furthermore, it has been shown that the ratio of carrier spatial frequency to envelope spatial frequency used to construct second-order stimuli, influences the probability of driving neurons with second-order stimuli. For example, applying various combinations of carrier and envelope spatial frequency demonstrated that different ratios of carrier spatial frequency to envelope spatial frequency (Zhou and Baker, 1994) instead of a fixed ratio (Albrecht and Hamilton, 1982) increases the probability of driving neurons with second-order stimuli. Furthermore, the spatial frequency tuning measurements of the
carrier have revealed that second-order responsive neurons prefer higher carrier spatial frequencies, and although the preferred ratio of carrier spatial frequency to envelope spatial frequency differs across neurons, this ratio tends to be around 8 (Li et al., 2014) (Figure 5). However, there is no fixed relationship between carrier and envelope spatial frequency and the preferred spatial frequency for the envelope is similar to that of LGs (Zhan and Baker, 2006; Li et al., 2014).

Another feature of second-order stimuli that has been investigated in detail is the carrier and envelope orientations. Previous studies have demonstrated that envelope orientation is the main factor of second-order responses (Mareschal and Baker, 1999a). However, further studies revealed the role of carrier orientation and spatial frequency in responsiveness to second-order stimuli. Indeed, neurons responsive to second-order stimuli are also selective to orientation and spatial frequency of carrier (e.g. Mareschal and Baker, 1999b). Theses findings led to further investigations which constructed second-order stimuli using different combinations of carrier and envelope orientations and also to find a bigger number of responsive neurons to second-order stimuli (Li and Baker, 2012; Li et al., 2014). Although there is no fixed relationship between the optimal orientation of carrier and that of envelope (Mareschal and Baker, 1999a; Li et al., 2014), second-order responsive neurons have been shown to be well tuned to carrier orientations. The finding of carrier orientation tuning in second-order responsive neurons rules out any source of nonlinear artifact in those neuron responses (Zhou and Baker, 1993; Li et al., 2014) (Error! Reference source not found.). Indeed, those artifacts which might potentially arise from display device or photoreceptors, cannot yield such carrier tuning, or even if they can, it would be the same in every neuron. Therefore, the observed carrier tunings in previous studies that were variable among neurons and distinct from that of envelope, provided some evidence to exclude any contribution of artifacts in second-order responses.
Cortical organization of responsive neurons to second-order stimuli

The cortical organization of responsive neurons to second-order stimuli also has been investigated with imaging and electrophysiological methods (Zhan and Baker, 2006; Li and Baker, 2012). Intrinsic optical imaging in area 18 of cats revealed similar orientation preference maps for first-order and second-order stimuli. In fact, LGs and illusory and contrast envelopes evoked the same orientation preference maps in area 18. The orientation preference of neuronal populations in response to first- and second-order stimuli was very similar, such that a majority of neurons showed <30 degrees difference in orientation preference. However, the fraction of responsive neurons to second-order was less than that to first-order stimuli and neural responses to second-order were weaker than those to first-order stimuli. In addition, the population responses to contrast and illusory contours were immune to different carrier orientations and therefore second-order responses were dependent on envelope orientation rather than carrier. The similarities of first- and second-order maps led to the idea of homogeneous distribution of second-order neurons across the cortical surface (Zhan and Baker, 2006). However, conflicting with this notion, another study investigating the functional organization of responsive neurons to second-order stimuli, in area 18 cats, argued against the homogenous distribution of second-order neurons across the cortical surface and instead, showed that responsive neurons to second-order are highly clustered and organized in a columnar manner (Li and Baker, 2012). Indeed, nearby neurons shared similar spatial frequency and to some extent similar orientation preference of carrier, and moreover, they formed a columnar organization along the cortex. While both spatial frequency and orientation maps of carrier showed clustered organization, carrier orientation maps were more scattered and therefore, this mini-cluster organization could be missed in intrinsic optical imaging (Zhan and Baker, 2006), which accesses only the superficial layers and has limited spatial resolution (Li and Baker, 2012).

Figure 6. There is no fixed relationship between carrier (red) and envelope (blue) orientations. Other conventions as Figure 5 (Adapted from Li et al., 2014).
Filter-Rectify-Filter model

Linear summation is the well-known underlying mechanism of first-order stimuli to elicit a response in simple and complex cells. **Figure 7** illustrates how a conventional LG can activate simple cell sub-regions if it is presented with an optimal orientation and spatial frequency. As this figure shows, to evoke maximally effective summation, the orientation and spacing of dark and light bars have to be aligned with the excitatory and inhibitory sub-regions of the simple cell. Although complex cells do not show such spatially segregated excitatory and inhibitory sub-regions, they are selective to different orientations, spatial frequencies and velocities. Hubel and Wiesel in 1962 (Hubel and Wiesel, 1962) suggested that complex cells sum up the responses from neighboring simple cell type neurons, which have similar preferences, but slightly different receptive filed positions. Therefore, complex cells’ orientation selectivity could be also explained by linear summation. However, a linear summation mechanism cannot account for second-order processing. As **Figure 8** shows while linear mechanism could account for luminance-modulated responses, CG responses could only be explained by a nonlinear mechanism. Although light and dark bars of an optimal LG can stimulate sufficiently the excitatory and inhibitory sub-regions of the receptive field of a simple cell and evoke a maximal response, the spatial frequency of the noise texture in CGs is too high to provide any net luminance changes within the neuron’s receptive field. In fact, light and dark elements of the carrier cancel out within individual sub-regions of the receptive

---

**Figure 7.** Luminance grating presented in an optimal spatial frequency and orientation evokes maximal neural responses (Adapted from Baker and Mareschal, 2001).

**Figure 8.** Linear mechanism cannot explain the second-order responses, unlike to first-order (Adapted from Baker and Mareschal, 2001).

**Potential model for second-order processing**
field and thus linear summation cannot account for contrast-modulated responses. Although the underlying mechanism of second-order processing has been intensively debated (e.g., Tanaka and Ohzawa, 2009; El-Shamayleh and Movshon, 2011; Hallum and Movshon, 2014; Li et al., 2014), human psychophysics (Landy and Graham, 2004) and electrophysiological studies in higher-order mammals (Movshon et al., 1978b; Issa et al., 2000) suggested a two-stage model of second-order processing (Zhan and Baker, 2006). This model consists of two linear filters and one non-linear rectifier (Figure 9). According to this filter-rectify-filter (FRF) model, neurons with small receptive fields would respond to the carrier, and their output is rectified and summed by the second and larger linear filter of low spatial frequency, providing orientation-selectivity to the envelope. However, it is currently unclear which brain structures serve as filters, in particular the first-stage filter. Considering the established spatial frequency and orientation tuning of carrier, the potential candidate for the first stage has to be selective for those features. V1 neurons might be the primary potential candidates to play the first-stage filter role in filter-rectify-filter model, as they are selective for spatial frequency and orientation of carrier (Mareschal and Baker, 1999a). However, recent studies suggested that LGN neurons could provide carrier-tuned input for V1 neurons, as they are spatial frequency and orientation selective to carrier (Rosenberg et al., 2010; Rosenberg and Issa, 2011). In any manner, first-stage information will be carried by V1 neurons, which constitute the subunits of V2 receptive fields, and thus the complex images could be perceived by V2 neurons (Li et al., 2014).

### Figure 9
Two-stage model suggested for second-order processing (Adapted from Baker and Mareschal, 2001).

**Surround suppression mechanism**

Although several studies have postulated the filter-rectify-filter model as the underlying mechanism of second-order responses, it is uncertain if such a mechanism exists in primates and other mammals (e.g., Tanaka and Ohzawa, 2009; El-Shamayleh and Movshon, 2011; Hallum and Movshon, 2014; Li et al., 2014). Indeed, another series of second-order studies proposed that responses to second-order stimuli could arise from other mechanisms, such as surround suppression (Tanaka and Ohzawa, 2009; El-Shamayleh and Movshon, 2011; Hallum and Movshon, 2014). Tanaka and Ohzawa suggested that the spatial organization of center-surround of neurons in area 17 of cats
is suitable to encode the orientation of higher-order contours. Based of neural responses to second-order stimuli, they reconstructed the spatial organization of V1 classical receptive fields and their surrounds and found that center-surround receptive fields of V1 neurons, generally, are organized elongated and parallel to each other. These findings led them to conclude that center-surround receptive fields act as a spatial filter and extract the higher-order borders by their filter shapes (Tanaka and Ohzawa, 2009).

**Mice as a popular model system for visual processing**

Over the past few years, interest in mouse visual cortex has considerably grown. Until now, the non-human primate has been the preferred animal model for studies of cortical visual processing because it has an elaborate visual system whose key properties, such as acuity or color vision, are mostly similar to those of humans. In many other areas of neuroscience, however, mice have played a prominent role (Boulanger, 2009; Crook and Housman, 2011; Mentis et al., 2011), since the mouse is currently the only mammal in which genetic engineering methods are routinely employed. Here, the advent of transgenic and knockout techniques has provided abundant possibilities to study normal and abnormal brain function and its relationship to brain structure. Because the availability of such powerful methods holds the potential to resolve long-standing questions of cortical visual processing (Liu et al., 2009b) interest in the mouse model has grown in the field of visual systems neuroscience. In studies of cortical visual processing, the use of a system as simple as the mouse cortex, which lacks both fine-scale spatial acuity and columnar maps such as those for orientation, promises to determine minimal mechanisms necessary for receptive field development and function (Niell and Stryker, 2008b). Remarkably, recent studies show that some key response properties of mouse visual neurons are actually quite comparable to those known from decades of study in higher mammals (Niell and Stryker, 2008b; Gao et al., 2010b; Van den Bergh et al., 2010). These properties include the sharp tuning for orientation, luminance-invariant contrast sensitivity and contrast-dependent temporal and spatial frequency sensitivity.

**Functional organization of mouse higher visual areas**

Although the mouse visual cortex is much more simple in both structure and function compared to primates and higher mammals, it still shares some basic similarities (Niell and Stryker, 2008a; Gao et al., 2010a; Van den Bergh et al., 2010) For example, anatomical studies have established around 10 separated, retinotopically organized areas in mouse visual cortex (Figure 10). These areas include the primary visual area, which is surrounded by nine other extrastriate areas, each containing a complete map of the entire visual hemifield (Wang et al., 2007b). Despite the ongoing debate on the functional division of rodent extrastriate areas, studies based on cytoarchitectonic and chemoarchitectonic markers and pathway tracing have categorized these areas into two parallel streams: the ventral stream which consists of LM, LI, POR, P and which is suggested to be responsible for object recognition; and the dorsal stream which consists of AL, PM, AM, A, RL and is proposed to be involved in object localization (Wang et al., 2011, Wang et al., 2012). Furthermore, distinct mouse extrastriate areas have been shown to have, to some extent, functional differences as well, when they are probed with first-order gratings (Van den Bergh et al., 2010; Andermann et al., 2011; Marshel et
For example, neurons in area AL, the gateway to the dorsal stream, prefers visual stimuli with higher temporal frequency and lower spatial frequency and therefore might be involved in object localization, while neurons in PM are driven better by high spatial frequency and low temporal frequency gratings and thus might have a role in detecting fine texture (Andermann et al., 2011; Marshel et al., 2011). Therefore, both anatomical and functional studies support the idea of existence of ventral and dorsal streams in mouse visual cortex, similar to those in higher mammals, but probably in a basic and simple manner.

Figure 10. Extrastriate visual areas in the primate and mouse (Adapted from Niell, 2011)

Can mice vision rely on more than luminance cues?
Although the number of studies exploring mouse vision has been growing in recent years, it is still not clear whether and how mice can extract meaningful representations from visual cues other than luminance. Studies comparing the basic properties of mouse visual areas (Van den Bergh et al., 2010; Andermann et al., 2011; Marshel et al., 2011) and examining the functional specificity of cortico-cortical projections (Glickfeld et al., 2013) used only first-order, LGs. Although using LGs offers several advantages such as direct comparison of the properties of extrastriate and V1 neurons and precise measurements of basic and discriminating features of ventral and dorsal streams (i.e. spatial frequency and temporal frequency), they may not drive higher-order neurons optimally. Also, in psychophysical studies where mice had to perform visual tasks (Andermann et al., 2010; Histed et al., 2012; Lee et al., 2012; Bennett et al., 2013), so far, only LGs have been employed. Therefore, it is not established yet whether and how the mouse visual system represents other visual cues than luminance and if mice can use those second-order stimuli to guide visual behavior.

Area LM might be a suitable area to investigate second-order processing
Among mouse extrastriate areas, LM has been thought to be homologous to area V2 in higher-order mammals, where selectivity for second-order contours has been found (Zhou and Baker, 1994; Leventhal et al., 1998; Mareschal and Baker, 1998a, b; Zhan and Baker, 2006; Song and Baker, 2007; Li et al., 2014). V2 and area LM are similar in several aspects: both areas share the representation of the vertical meridian with V1 (Coogan and Burkhalter, 1993; Wang et al., 2007b), both areas are primary targets of V1 projections (Wang et al., 2011) and both contains neurons with ventral and dorsal
stream properties (Wang et al., 2012). Another reason that makes LM a suitable area to study second-order processing in mouse is the potential role of this area in object recognition, as anatomical studies identified area LM as the gateway of the mouse ventral stream (Wang et al., 2011). Furthermore, one study in rodents showed that LM lesions could disrupt some aspects of object recognition (Sacco and Sacchetti, 2010). Taken together, the established properties of area LM make this area as an interesting, putative region involved in second-order processing.

**Can mice use second-order stimuli to guide visual perception in a cue-invariant way?**

In this study I investigated whether mice can use second-order stimuli to guide visual perception in a cue-invariant way and examined potential cue-invariant representations of stimulus orientation in two areas of mouse visual cortex. Using a classical condition task, I tested whether orientation discrimination could be achieved invariant to defining cue of visual stimuli. More precisely, I asked if mice can generalize orientation discrimination from first-order LGs to second-order, CGs. I found that mice are capable of generalizing orientation discrimination learning from familiar cue conditions to novel cue conditions. Then I performed extracellular recordings of neural responses to LG and CG stimuli in mouse areas V1 and LM, and compared the orientation tuning curves to these two types of gratings. Although in both areas the responses to CGs compared to those to LGs were weaker and less selective, neurons showed broadly similar preferences for the two types of stimuli. I conclude that mice are able to use second-order stimuli to guide visual perception, albeit in a basic form.
Material & Methods
All procedures complied with the European Communities Council Directive 2010/63/EC and the German Law for Protection of Animals, and were approved by local authorities, following appropriate ethics review.

Luminance-modulated and contrast-modulated gratings
Visual stimuli were presented using custom software (EXPO; https://sites.google.com/a/nyu.edu/expo/home) on a calibrated liquid crystal display (LCD) monitor (Samsung 2233RZ; mean luminance 50 cd/m²), placed 25 cm in front of the animal’s eyes. To correct luminance nonlinearities of the display I used an inverse gamma lookup table obtained regularly by calibration with a photometer.

For both behavioral and electrophysiological experiments, I used a common set of luminance-modulated gratings (LGs) and contrast-modulated gratings (CGs), obtained by multiplying a contrast envelope with a noise carrier. I defined a two-dimensional moving sine wave grating as: $S(x,y,t) = \sin(2\pi f (\sin(\theta) x + \cos(\theta) y) + 2\pi vt)$, where $f$ is the spatial frequency, $\theta$ is the orientation, and $v$ is the temporal frequency. LGs were then generated as: $LG(x,y,t) = l_0 + l_0 CG(x,y,t)$, and CGs were generated as $CG(x,y,t) = l_0 + l_0 (S(x,y,t) + 1)/2 * N(x,y)$, where $l_0$ is the mean luminance, $c$ is the contrast and $N(x,y)$ is the static noise carrier with a spatial frequency spectrum that dropped off as $A(f_n)\sim 1/(f_n + f_c)$. I created two types of CG stimuli that differed in the distribution of Fourier energy of the carrier. For CGs with low-frequency noise, I set $f_c$ to 0.05 cycles/degree and imposed a high frequency cutoff at 0.12 cycles/degree; for CGs with high frequency noise, I used an $f_c = 0$ cycles/degree and a low frequency cutoff at 0.12 cycles/degree. For recordings from area V1, the LGs and the envelope of the CGs had a spatial frequency $f$ of 0.05 cycles/degree and a temporal frequency $v$ of 1.5 Hz. To optimize stimulus parameters for the preferences of area LM, I conducted the LM recordings with $f$ of 0.028 cycles/degree and $v$ of 1.8 Hz (Marshel et al., 2011). Contrast $c$ was set to 1, except for experiments where I matched the root-mean-square (RMS) contrast between first- and second-order gratings, in which case $c$ was 0.335 for the LGs. The seed for generating the random Gaussian noise texture was varied across experimental sessions.

Analysis of visual stimuli
Following An et al. (2014), I performed a spectral power analysis of the LGs and CGs. To reveal the difference in power between two orthogonal orientations, I first calculated $D(\omega_x, \omega_y, \tau) = P_o(\omega_x, \omega_y, \tau) - P_o(\omega_x, \omega_y, \tau)$, where the power $P$ is the squared amplitude of the 3-dimensional Fourier transform of the drifting grating. To illustrate the difference in power as a function of space, $D$ was further integrated to result in $I(\omega_x, \omega_y) = \int D(\omega_x, \omega_y, \tau)d\tau$. To illustrate the difference in power as a function of orientation, $I(\omega_x, \omega_y)$ was transformed into $I(\rho, \theta)$ and was further integrated to result in $O(\theta) = \int I(\rho, \theta)dp$. To assess the spatial frequency and temporal frequency content of the absolute differential power, I calculated $S(\rho) = \int |I(\rho, \theta)|dp$ and $T(\tau) = \iint |P_o(\omega_x, \omega_y, \tau) - P_o(\omega_x, \omega_y, \tau)|d\omega_x d\omega_y$. 

18
Surgical preparation for head-fixed visual behavior

Behavioral experiments were performed in 2-5 months old C57BL/6J mice of either sex (2 males, 4 females). General anesthesia was induced by 5% isoflurane and maintained during surgery at 1-2%. Buprenorphine (0.1 mg/kg, sc) was used for analgesia and atropine (0.3 mg/g, sc) to reduce bronchial secretions. Animal temperature was kept at 37°C. A custom-designed head post was mounted to the skull using dental cement (Tetric EvoFlow, Ivoclar Vivadent). Mice were implanted with two miniature screws over the cerebellum (#00-96X 1/16, Bilaney), serving as reference and ground with extracellular recordings. The skull over the target area was marked and sealed with KwikKast (WPI). For 3 days after surgery, mice were injected by antibiotics (Baytril, 5mg/kg, sc) and longer lasting analgesics (Carprofen, 5mg/kg, sc). After recovery, mice were gradually habituated to being head-fixed and placed on an air-suspended Styrofoam ball (Holscher et al., 2005; Dombeck et al., 2007) (Figure 1a). A spout connected to a lick sensor was used to measure licks and deliver fluid rewards (Schwarz et al., 2010).

Orientation discrimination task and analysis of behavioral data

After habituation to the setup, mice were placed on a water restriction regimen. Throughout all training phases, the animals’ daily weight and fluid consumption were monitored and recorded, and the animals were checked for signs of potential dehydration (Guo et al., 2014). Following Guo et al. (2014), in a first phase, daily access to water was systematically reduced until the animal reached a target weight of ~85% of its initial weight. After the weight had stabilized, training in the behavioral task started and mice received most of their water during performance in the behavioral apparatus.

Using classical conditioning, mice were trained to associate the orientation of the visual stimulus with a water reward. In each session, mice were presented with 120 trials of either 45 deg or 315 deg gratings, drifting for a duration of 3 s behind an aperture of 32 deg diameter. The presentation of the 315 deg grating was automatically followed by a fluid reward of 5-7 µl, the orthogonal grating was never rewarded. Stimulus presentations were separated by an interstimulus interval of 15 s added to a random delay drawn from an exponential distribution with a mean of 15 s. Drawing onset times from an exponential distributions yields a flat hazard rate, ensuring that animals cannot predict the time point of reward delivery.

To evaluate orientation discrimination performance, I focused on licks in anticipation of fluid reward. Following Gallistel et al. (2004), I computed, separately for each stimulus orientation, a lick index $LI = (\text{licks}_{\text{stimulus}} - \text{licks}_{\text{baseline}})/(\text{licks}_{\text{stimulus}} + \text{licks}_{\text{baseline}})$, where licks_{stimulus} is the number of licks during the last 1 second of stimulus presentation and licks_{baseline} is the number of licks during the 1 second before stimulus presentation. To identify learning I analyzed the cumulative records of $LI$, where changes in slope correspond to changes in the level of performance. For instance, a positive slope of the cumulative $LI$ corresponds to increased licking during the stimulus compared to the baseline period, and indicates that the animal anticipates reward after seeing any of the stimuli. A positive slope of the difference between cumulative $LIs$ indicates that the animal licks more strongly during the rewarded than during the unrewarded grating, and shows that the animal has learned to discriminate grating orientations. To assign trials to different stages of orientation discrimination learning, I determined significant
changes in the slope of the difference of the cumulative $LI$ by using a change point analysis (Gallistel et al., 2001). To quantify discrimination performance across the different stages of learning, I performed an ideal observer analysis on the distributions of $LIs$ for the two orientations.

I tested the mice in several conditions. I always started training using $LGs$ until the animal reached stable and reliable orientation discrimination performance. Then, I replaced $LGs$ with $CGs$ with low-frequency noise carriers to test the generalization of orientation discrimination to second-order gratings. To test for influences of stimulus contrast, I switched back to $LGs$, but with contrast reduced to 0.335, such that it matched the RMS contrast of $CGs$. After that, 2 animals were additionally tested with $CGs$ with high-frequency noise carriers.

**Surgical preparation for electrophysiological recordings during wakefulness**
Electrophysiological recordings were performed in 2-5 months old C57BL/6J mice of either sex (4 males, 8 females). Surgical procedures were identical to those for behavioral experiments. After recovery and habituation to the setup, mice underwent a second surgical procedure under general anesthesia (1-2% Isoflurane, ~15 min), in which a craniotomy was performed, which was sealed with KwikKast until the recording session. To avoid potential effects of anesthesia, recordings were never performed on the same day of the craniotomy.

**Surgical preparation for electrophysiological recordings during anesthesia**
For electrophysiological recordings under anesthesia, I used 4 C57BL/6J mice. Anesthesia was induced by 5% isoflurane and maintained during surgery by a combination of urethane (375 mg/kg, ip), chlorprothixene hydrochloride (2 mg/kg, ip) and isoflurane (1–2%). I used Buprenorphine (0.1 mg/kg, sc) analgesia and atropine (0.3 mg/g, sc) to reduce bronchial secretions. Animal temperature was kept at 37°C. A custom-designed head post was mounted to the skull using dental cement (Tetrik EvoFlow, Ivoclar Vivadent). As reference, I used a wire placed into the cerebellum; as ground, a wire placed under the exposed skin. During recordings, isoflurane was kept at minimal to reduce unwanted side effects on visual responses (Vaiciulionaita et al., 2013). Depending on the level of anesthesia as assessed by the breathing rate and presence or absence of the pinch toe reflex, additional doses of chlorprothixene hydrochloride (2mg/kg, ip) were injected (approximately every 3 h). Recordings typically lasted for 6 h.

Recordings from V1 were obtained through a craniotomy (<1.5 mm) located 3 mm lateral to the midline and 1.1 mm in front of the anterior margin of the transverse sinus. Recordings from LM were obtained from 4 mm lateral to the midline and 1.4 mm in front of the anterior margin of the transverse sinus “Wang, 2011”. Recordings from V1 and LM were performed in separate sessions using 32-channel silicon probes in a 4-shank configuration (Buzsaki32-A32, Neuronexus; 200 inter-shank spacing). Extracellular signals were recorded at 30 kHz (Blackrock Microsystems). Online estimates of tuning properties relied on high-pass filtered signals crossing a fixed threshold (typically 4.5–6.5 SDs).
Visual stimuli for electrophysiological recordings
To estimate receptive field (RF) position, I mapped ON and OFF subfields of RFs using a sparse noise stimulus (Liu et al., 2009a). This stimulus consisted of white or black squares (4° diameter) briefly flashed (150 or 200 ms) on a square grid (40 or 60° diameter). Subsequent stimuli were centered on the online estimates of the average RF maps for each shank of a 32-channel silicon probe (see “Analysis of electrophysiological data” and Figure 1b). To measure orientation tuning curves, I interleaved in pseudorandom order LGs and CGs moving for a duration of 2 s in 8 different directions.

Analysis of electrophysiological data

Unit extraction and spike sorting
Recordings from V1 were obtained through a craniotomy (~1.5 x 1.5 mm) located 3 mm lateral to the midline and 1.1 mm in front of the anterior margin of the transverse sinus. Recordings from LM were obtained from 4 mm lateral to the midline and 1.4 mm in front of the anterior margin of the transverse sinus (Wang et al., 2011). Recordings from V1 and LM were performed in separate sessions using 32-channel silicon probes in a 4-shank configuration (Buzsaki32-A32, Neuronexus; 200 μm inter-shank spacing, Figure 1b). Extracellular signals were recorded at 30 kHz (Blackrock microsystems). Online estimates of tuning properties relied on high-pass filtered signals crossing a fixed threshold (typically 4.5–6.5 SDs).

Wideband extracellular signals were digitized at 30 kHz (Blackrock microsystems) and analyzed using the NDManager software suite (Hazan et al., 2006). The 8 channels on each shank were treated as an “octrode”. Using a robust spike detection threshold (Quiroga et al., 2004) set to 6 standard deviations of the background noise, spike-waves were extracted for each “octrode” from the high-pass filtered continuous signal. The first 3 principal components of each channel were used for automatic clustering with KlustaKwik, followed by manual refinement of clusters (Hazan et al., 2006). This yielded high-quality single unit activity as evident from distinct spike wave shapes and a clear refractory period in the autocorrelogram (Figure 1d). For analysis of retinotopy, I used the envelope of the multiunit activity (MUAe, van der Togt et al., 2005), averaged across all channels in each shank.

Analysis of tuning
To determine RF maps for single unit spiking activity, I fitted ON and OFF subfields separately with a two-dimensional Gaussian (Liu et al., 2009a):

\[ f(x, y) = B + \frac{A}{2\pi ab} \exp \left( -\frac{x'^2}{2a^2} - \frac{y'^2}{2b^2} \right) \]

where \( A \) is the maximum amplitude, \( B \) is the baseline response, \( a \) and \( b \) are half-axes of the ellipse, and \( x' \) and \( y' \) are transformations of the stimulus coordinates \( x \) and \( y \), taking into account the angle \( \theta \) and the coordinates of the center \( (x_c, y_c) \) of the ellipse. To quantify the progression of RF location in the recorded visual area, I constructed maps of z-scored MUAe activity, averaged between 0 – 0.35 s after stimulus onset. If these maps had a sufficient signal/noise ratio (standard deviation > 0.035), I computed the average RF coordinates from the peak of the MUAe activity for each shank. Sessions with ambiguous maps were discarded from all further analyses.

Orientation tuning curves were fitted with a sum of two Gaussians with peaks 180 deg apart, which could have different amplitudes but equal width and a constant baseline. To quantify orientation selectivity, I computed d-prime (Berens et al., 2008) defined as
\[ d' = \frac{\mu_{\text{pref}} - \mu_{\text{ortho}}}{\sigma}, \text{ where } \sigma = \sqrt{\left(\sigma_{\text{pref}}^2 + \sigma_{\text{ortho}}^2\right)/2}. \]

In contrast to the more commonly used OSI (Niell and Stryker, 2008a), this index has the advantage to not only consider modulation depth but also the variability of responses. I performed all our analyses also on OSI (both with and without spontaneous activity removed), and circular variance (Ringach et al., 2002); results obtained with these alternative measures were qualitatively similar. I only considered neurons which passed 3 selection criteria applied to the responses to LGs: (1) an average firing rate of at least 1 spike/s to at least one orientation, (2) average responses to at least 2 orientations differing from the response to the mean-luminance gray screen by at least 2.58 times the standard error of the mean, and (3) explained variance of the Gaussian fit of at least 70%.

**Comparison of responses to luminance-modulated and contrast-modulated gratings**

On the population of selected neurons, I performed a log-linear analysis to statistically assess the proportion of neurons responsive to both LGs and CGs vs. LGs only. To model the observed counts, I fitted a generalized linear model (GLM) with a Poisson link function considering the factors responsiveness (LGs and CGs vs. LGs only) and area (V1 vs. LM). I report all significant interactions with the factor responsiveness. In addition, I included the factor noise (low-frequency vs. high-frequency) to assess differences in responsiveness across experiments with different noise textures.

To investigate differences in firing rates and orientation tuning in response to LGs and CGs I performed an analysis of variance (ANOVA) with the within-subject factor stimulus (LGs vs. CGs), and the between-subject factors area (V1 vs. LM). To appropriately visualize the results, I show the mean and standard error of the pairwise differences (Franz and Loftus, 2012). To compare the difference in responses to CGs vs. LGs, for LGs with full contrast and matched RMS contrast, I performed an ANOVA with the within-subject factor stimulus (LGs vs. CGs), and the between-subject factors contrast (full vs. matched LG) and area (V1 vs. LM). All post-hoc pairwise contrasts were corrected for multiple comparisons (multcomp package in R, (R Development Core Team, 2015).

To relate the preferred orientations in response to LGs and CGs I used the circular version of the Pearson’s product moment correlation as described by (Jammalamadaka and Sengupta, 2001) and implemented in Matlab by (Berens, 2009).

**Histology**

For post-mortem histological reconstruction of recording sites (Figure 1c) I coated each shank of the electrode alternating between a red-shifted fluorescent lipophilic tracer (DiD; D7757, Invitrogen) and an orange fluorescent lipophilic dye (DiI; D282, Invitrogen). After recordings, mice were transcardially perfused under pentobarbital sodium anesthesia (200 mg/kg) with 0.2 M sodium phosphate buffer (PBS), followed by 4% paraformaldehyde in PBS. Brains were postfixed for 24 h at 4°C and then rinsed 3 times with 1x PBS. Brains were sliced (40 μm) using a vibratome (Microm HM 650 V-Thermo Scientific) and mounted on glass slides with Vectashield DAPI (Vector Laboratories), and coverslipped. Slides were inspected for the presence of the tracers using a Zeiss Imager.Z1m fluorescent microscope.
Figure 1. Behavior setup, recordings, and isolated single neurons. 

- **a**, Setup for head-fixed behavior with air-cushioned spherical treadmill and lick-sensor.
- **b**, Schematic of the four-shank silicon probe.
- **c**, Coronal section. The four shanks of the electrode were stained in alternating fashion with Dil (yellow) and DiD (red). Blue represents DAPI. Scale bar, 400 μm.
- **d**, Average spike-waveforms and autocorrelograms of two example single units recorded from area V1. Units 144–5.x.31 and 144–4.x.5.


Results

Stimuli
To discriminate the global, first-order and local, second-order features of CGs, I performed a spectral analysis of the stimuli used in this study. Figure 1 shows the spectral analysis performed on the grating orientations used for the behavioral experiments. LGs (Figure 1a1) contain power at a single spatial frequency and orientation (Figure 1a2) and the peaks of their differential power in the orientation domain correspond to grating orientations, i.e. at −45 and 45 degrees (Figure 1a3, a4). The absolute differential power in the spatial frequency domain peaks at 0.05 cycles/degree (Figure 1a5) and the absolute differential power in the temporal frequency domain peaks at 1.5 Hz (Figure 1a6), which correspond to the grating spatial frequency and temporal frequency, respectively. For CGs with low-frequency noise carrier (Figure 1b1, b2), the differential power peaks at −45 and 45 degrees, which correspond to the orientation of envelope (Figure 1b3, b4). This is due to an inhomogeneity in the distribution of local luminance of the noise carrier, induced by the envelope. However, the magnitude of this differential power is considerably different from that for LGs, it is smaller (25% of that for LGs) and more broadly distributed across orientations. In the spatial frequency domain also, CGs with low-frequency noise show a broad differential power distribution between 0.02 and 0.15 cycles/degree (Figure 1b5), while in the temporal frequency domain, absolute differential power peaks clearly at 1.5 HZ, corresponding to the drift rate of the envelope (Figure 1b6). In comparison, the spectral analysis on CGs with high-frequency noise carrier (Figure 1c1, c2) revealed only little differential power across spatial frequencies (Figure 1c3), as no clear peaks and troughs of differential power were observed around the orientations of the LGs (Figure 1c4). For this type of CGs, the absolute differential power in the spatial frequency domain is small for values lower than 0.08 cycles/degree, highest at 0.16 cycles/degree, and diminishes toward higher spatial frequencies (Figure 1c5); the absolute differential power in the temporal frequency domain again peaks at 1.5 HZ, which is the drift rate of the envelope (Figure 1c6).

Together, this spectral analysis reveals the difference between the two types of CGs in activating the first-order luminance-sensitive mechanism in the mouse visual system: while the distortion signals in CGs with low-frequency noise carriers could lead to decoding stimulus orientation via first-order mechanisms, it seems unlikely that CGs with high-frequency noise carrier invoke the same mechanism. Indeed, as explained above, for CGs with high-frequency noise carrier, there are no clear peaks and troughs around the orientation of the LGs in the differential orientation signal across spatial frequencies and the spatial frequencies of the noise carrier are concentrated beyond the optimal values for mouse V1 (0.045 cycles/degree) and LM (0.028 cycles/degree) neurons (Marshel et al., 2011).
Orientation discrimination learning for luminance-modulated gratings

I wanted to test whether mice can use second-order stimuli to guide visual perception in a cue-invariant way. In order to address this question, first I trained head-fixed mice in a classical conditioning paradigm to discriminate different orientations of LGs (Figure 2). I presented two orthogonal orientations and paired one of these orientations with a fluid reward. Next, I measured discrimination performance with analyzing the orientation-specific licking in anticipation of reward. At the beginning of training, animals licked constantly, regardless of stimulus presentation and increased lick rates to pick up the reward (Figure 2a). However, when animals learned the task, a high rate of licks was observed during the presentation of the rewarded orientation (Figure 2b). To evaluate licking, I computed, for each orientation separately, the cumulative sum of a lick index (LI), defined as the difference in the number of licks during the last 1 s of stimulus presentation and the 1 s before stimulus presentation, divided by their sum (Figure 2c). Then, I calculated the difference between cumulative LIs to the rewarded and unrewarded orientations and identified significant change points, which are related to significant changes in orientation discrimination performance (Figure 2d). Later, I performed an ideal observer analysis on the animal’s LI, which indicated that orientations could not be decoded before (area under the receiver operating characteristic [AUROC] = 0.52, 95% CI = 0.50– 0.55), but after learning (AUROC = 0.78, 95% CI = 0.75– 0.80). Finally, I summarized the performance across all mice and found similar results, with average performance increasing from $0.51 \pm 0.008$ (SEM) to $0.80 \pm 0.023$. 

Figure 1. Analysis of visual stimuli. a, LGs (32 degrees wide) of orthogonal orientations ($a_1$) and their power spectra ($a_2$). Differential power in the Fourier plane ($a_3$) and orientation domain ($a_4$). Spatial frequency ($a_5$) and temporal frequency content ($a_6$) of the absolute differential power distribution. b, Same as a, for the CGs with low-frequency noise. c, Same as a, for the CGs with high-frequency noise. Red circles and lines indicate a spatial frequency of 0.1 cycles/degree.
**Figure 2.** Orientation discrimination for LGs. **a.** Example behavioral session, before learning. Top, Licks to the rewarded orientation. Middle, Licks to the unrewarded orientation. Bottom, Trial-averaged lick density. Shaded regions represent mean SEM. Gratings indicate stimulus onset. Black vertical lines indicate stimulus offset. Black horizontal lines indicate baseline and stimulus period used for computing the LI. Session 5. **b.** Same as **a**, after learning. Session 21. **c.** Cumulative LI as a function of trial number for rewarded (blue) and unrewarded orientations (red). **d.** Difference of cumulative LIs between the two orientations. Dots indicate significant change points used for assigning sessions to training stages. Trials before the first change point were assigned to the pre-learning stage (dashed vertical line). Trials after the last change point (solid vertical line) were assigned to the post-learning stage. **e.** ROC analysis based on distributions of LIs from the two conditions in the pre-learning stage (dashed curve) and post-learning stage (solid curve). **a-e.** Example mouse 278. **f.** AUROC across mice (N=6). Crosses represent 95% CIs.

**Cue-invariant generalization of orientation discrimination**

Next, I tested whether mice can generalize the learned orientation discrimination from LGs to CGs, in which orientation was characterized by changes in contrast rather than luminance (**Figure 1b**). After mice had successfully learned the orientation discrimination task and reached a stable and reliable performance, I switched only the stimulus from LGs to CGs and kept all other aspects of the task identical. I replaced LGs, first, with low-frequency noise carriers CGs because this would help to transfer of learning easier. Although the global appearances of LGs and CGs with low-frequency noise carrier are different, mice could, in principle, discriminate the two orientations of the low-frequency CGs by relying on first-order mechanisms similar to those optimal for the learned task with LGs. The positive slopes of the cumulative LIs for rewarded and unrewarded stimuli indicate that mice could distinguish the CGs from the mean-luminance gray background. I also found that all tested animals, with the exception of one mouse, could judge the orientations of CGs with low-frequency noise carrier, as evident from the steeper increase of LI for the rewarded compared with the unrewarded condition (**Figure 3a**). Furthermore, no significant changes were found in
the slope of the differential cumulative $LI$ across trials, which means mice did not need to learn the orientation discrimination task from scratch but instead could readily generalize from $LG$s to $CG$s (Figure 3b). The fact that the slope of the differential cumulative $LI$ does not significantly change across trials also indicates that even extensive training of almost 2000 trials did not improve the performance for $CG$s. Overall, mice showed a considerably lower performance for $CG$s with low-frequency noise (AUROC = 0.63, 95% CI = 0.62–0.65; Figure 3c) than for the $LG$s and all tested mice showed the similar results (mean AUROC = 0.58 ± 0.017 SEM; Figure 3d).

I wondered whether the lower RMS contrast of $CG$s compared with $LG$s could account for the overall lower performance for $CG$s. To test this hypothesis, I performed control experiments in which the RMS contrast of $LG$s was lowered to match that of $CG$s and trained mice with these RMS contrast matched $LG$s in the next phase. Mice showed a good performance during the orientation discrimination task for $LG$s matched in RMS contrast (mean AUROC = 0.83 ± 0.03 SEM; Figure 3e–g). Indeed, summarizing the performances across all tested mice revealed similar results for both levels of contrast ($p = 0.23$, paired $t$ test; Figure 3h).

Finally, I probed mice with the other type of $CG$s, for which I imposed a low-frequency cutoff on the noise carrier. This type of $CG$s did not contain any energy at the spatial frequency of the learned $LG$s and little differential power across spatial frequencies at the learned orientations (Figure 1c). Again, mice could distinguish $CG$s with high-frequency noise carrier from the mean-luminance gray background, evident in the positive slope of cumulative $LIs$ across trials (Figure 3i). Importantly, the mouse could also tell the two grating orientations apart, which is revealed by the increase in the difference of cumulative $LIs$ (Figure 3j); however, its performance for this type of $CG$ was again considerably lower compared with that for $LG$s (AUROC = 0.58, 95% CI = 0.56 – 0.60; Figure 3k). Similar results were obtained when I tested a second mouse (AUROC = 0.56, 95% = CI 0.53 – 0.60; Figure 3l). Together, these results reveal the ability of mice in using second-order stimuli to guide visual perception.
Identification of areas V1 and LM based on mirrored retinotopic representation of azimuth

To investigate potential neural correlates of observed cue-invariant generalization of orientation discrimination in behavioral experiments, I performed extracellular recordings from areas V1 and LM. It has been established that area LM in mouse shares the vertical meridian with V1 (Coogan and Burkhalter, 1993; Wang and Burkhalter, 2007) and therefore is thought to be homologous to area V2 in higher-order mammals, where neurons have been indicated to be selective for second-order contours (Zhou and Baker, 1994; Leventhal et al., 1998; Mareschal and Baker, 1998a, b; Zhan and Baker, 2006; Song and Baker, 2007; Li et al., 2014). In addition, area LM is the primary target of V1 projections (Wang et al., 2012) and prefers lower spatial frequencies than V1 (Marshel et al., 2011).

The mirrored progression of retinotopy along the azimuth in the two areas provided the advantage to verify that my recording sites were indeed in areas V1 and LM. For recordings, I used a 4-shank silicon probe (See “Material and methods” Fig. 1b,c) spanning a large range of azimuths in either area and mapped RFs by presenting a sparse-noise stimulus (Figure 4). In order to obtain RF maps individually for each neuron, two-dimensional Gaussians were fitted to the maps of average firing rates (Figure 4a,d, top), separately for ON and OFF stimuli (Figure 4a,d, middle). Finally, I
considered only RFs with well-fit Gaussian profiles and extracted the average azimuth and elevation per shank for those RFs (Figure 4b,e, gray). Considering the fact that RF locations in mouse cortex are considerably scattered (Smith and Hausser, 2010; Bonin et al., 2011), I also analyzed the MUAe (Super and Roelfsema, 2005) for which I determined the peak RF coordinates for each electrode shank (Figure 4a,d, bottom, b,e, black). Similar to previous results on the retinotopy of mouse visual areas (Schuett et al., 2002; Wang and Burkhalter, 2007), I found that in V1 area, going from the most medial to the most lateral electrode shank, the azimuth of RF centers shifted from more peripheral to more central (Figure 4c); whereas, conversely, in LM recordings, the azimuth of RF centers moved from more central to more peripheral (Figure 4f).

Responses to contrast-modulated gratings are weaker and less selective
After verifying the identity of the targeted area, I presented visual stimuli on the center of mapped multiunit RFs and measured the orientation tuning curves of neurons in response to LGs and CGs. Then, I compared the neural responses to LGs and CGs across areas V1 and LM. Consistent with previous studies in higher-order mammals (Albright, 1992; Zhou and Baker, 1994; Mareschal and Baker, 1998a; Zhan and Baker, 2006; Li et al., 2014), I found that a considerable percentage of neurons did not respond to CGs, despite significant responses to LGs (Figure 5).
Responses to low-frequency noise contrast-modulated gratings

I interleaved LGs and CGs with low-frequency noise carrier in 8 different orientations and compared the responses orientation tuning curves in response to LGs and CGs with low-frequency noise, across areas V1 and LM. Among all recorded neurons, only 72% of luminance-responsive neurons also responded to the CGs with low-frequency noise (Figure 6). The contribution of neurons, which were responsive not only to LGs but also to CGs with low-frequency noise as well, was higher in V1 (77%, 178 of 230 recorded neurons) than in LM (61%, 69 of 114 recorded neurons, \( p < 0.001 \), two-way interaction, log-linear analysis; Figure 6). Further analysis was performed only on those neurons with a significant response to both types of stimuli.

Among those neurons with significant responses to both LGs and CGs, example neurons in both V1 (Figure 7a) and area LM (Figure 7b) showed lower peak firing rates in response to CGs with low-frequency noise than to LGs. In the population of neurons, a similar pattern of reduced peak firing rates to CGs was observed. Peak responses across both V1 and area LM decreased by 27.8 ± 3.2%, from 11.2 spikes/s in response to LGs to 8.1 spikes/s in response to CGs (ANOVA, main effect, \( p < 10^{-16} \); Figure 7c,d). This reduction was less pronounced in area V1 (25.9 ± 3.1%) than in area LM (31.8 ± 7.3%; interaction, \( p < 0.065 \)). I also found that the responses to the orthogonal orientation for CGs versus LGs were similar in area V1 (mean change 1.6 ± 0.5%), while LM neurons, more consistently, showed a reduction in their CGs responses (mean change −18.6 ± 6.3%; interaction \( p = 0.0002 \); Figure 7e,f).

This pattern of changes in the responsiveness was reflected in the orientation selectivity, both in the example cells (Figure 7a,b) and in the population (Figure 7g,h): neurons in both areas showed poorer orientation selectivity in response to CGs with low-frequency noise carriers compared to LGs. To measure orientation selectivity, I computed \( d' \), which does not only take into account the difference between responses to preferred and orthogonal orientations but also the variability of responses (Berens, 2009). Comparing orientation selectivity across V1 and area LM revealed, overall, a lower \( d' \) for LM than for V1 (main effect, \( p = 0.003 \)). More importantly, \( d' \) was lower for
CGs than LGs (main effect, p < 10^{-16}), and this reduction was more prominent for V1 (60.8 ± 4.5% from 1.39 for LGs to 0.54 for CGs) than LM (52.0 ± 6.9% from 0.99 for LGs to 0.47 for CGs; interaction, p = 0.003). Therefore, responses to CGs with low-frequency noise carriers compared with LGs in mouse visual cortex are weaker and orientation selectivity was poorer.

Figure 6. Distribution of responsive and unresponsive neurons to CGs with low-frequency noise carrier in V1 (N = 230) and area LM (N = 114).

Figure 7. Responses to CGs with low-frequency noise carriers in mouse visual cortex. a, Responses to LGs (black) and CGs (red) of an example neuron from area V1. Top, Raster plots. Bottom, Tuning curves of the same example neuron (Unit 221-1-7.61). b, Same, for an LM example neuron (Unit 245-4-6.35). c, d, Response to the preferred orientation for CGs versus LGs in the population of responsive neurons recorded from V1, N = 178 (c) and LM, N = 69 (d). Insets, Mean pairwise differences and their SEM (Franz and Loftus, 2012). Stars represent significance of post hoc comparisons. e, f, Same as c, d, for response to orthogonal orientation. g, h, Same as c, d, for d'. Red crosses represent means. Conventions as in Figure 5.

Responses to RMS matched contrast luminance-modulated gratings

To test whether the weaker responses and broader orientation tuning for CG than LG responses could be explained by the lower RMS contrast of CG gratings, I performed control experiments (Figure 9). In these experiments, the RMS contrast of LGs was matched with that of CGs, and responses to this type of LGs were measured. In V1 area only 71% of responsive neurons to LGs with matched RMS contrast also responded to
the CGs with low-frequency noise, this fraction was 76% in area LM (Figure 8). Comparing responses to preferred orientation between CGs and LGs matched in RMS contrast revealed again weaker responses to CGs than LGs in both areas (7.3 spikes/s vs. 8.8 spikes/s, main effect, $p < 10^{-5}$; Figure 9), however, this reduction in responsiveness was less pronounced compared to conditions with full contrast LGs (27.8 ± 3.2%; compare Figure 7c,d and Figure 9c,d; interaction, $p = 0.009$). I also found that in area V1, CGs evoked stronger responses to orthogonal orientations than LGs (4.41 sp/s vs 3.65 sp/s, 20.9 ± 4.2%), however this was not the case for area LM, i.e. the responses to orthogonal orientations did not differ significantly between LGs and CGs (8.5 sp/s vs 8.7 sp/s, 2.1 ± 4.7%; interaction, $p < 0.01$). Overall $d'$ was again lower for area LM (0.62 ± 0.05) than V1 (0.86 ± 0.06; main effect, $p = 0.026$), even with matched RMS contrast. Similar to our results with full-contrast LGs, $d'$ dropped considerably between LGs matched in RMS contrast and CGs (66.2 ± 5.5%, from 1.2 for LGs to 0.4 for CGs, main effect, $p < 10^{-16}$), while $d'$ was not significantly different between responses to full-contrast and reduced-contrast LGs ($d'$ full = 1.3, $d'$ matched = 1.2; compare Figure 7g,h and Figure 9g,h; two-sample t test, $p = 0.4$), which is probably reminiscent of the well-known phenomenon of contrast invariance of orientation tuning (Movshon et al., 1978; Albrecht and Hamilton, 1982; Sclar and Freeman, 1982). Interestingly, reductions in orientation selectivity ($d'$) between CGs and LGs with matched RMS contrast were more prominent in area V1 (68.3 ± 6.0%, from 1.3 to 0.41) than in area LM (56.3 ± 11.5%, from 0.86 to 0.38; interaction, $p = 0.008$).

Together, the reduced RMS contrast of CGs might explain the weaker responses to CGs but cannot account for the poorer orientation selectivity for CGs. Instead, the broader distribution of orientation energy in CGs compared with LGs might partially be responsible for poorer orientation selectivity for CGs (Figure 1,b).

**Figure 8.** Distribution of responsive and unresponsive neurons to CGs with low-frequency noise carrier in RMS-matched contrast LGs experiments, in area V1 ($N = 105$) and LM ($N = 59$).
Responses to high-frequency noise contrast-modulated gratings

Given that first-order, luminance-sensitive mechanisms in V1 and LM might contribute to the responses to CGs with low-frequency noise, I performed additional experiments to test the level of this contribution. Indeed, my findings of weaker responses to preferred orientations of CGs, of stronger responses to orthogonal orientations and of poorer orientation tuning compared with LGs are consistent with a potential activation of luminance-sensitive mechanisms. To decrease the local orientation-biased luminance fluctuations of the CGs in the preferred range of spatial frequency for V1 and LM neurons, I imposed a low-frequency cutoff on the noise carrier such that the noise carrier’s spatial frequency distribution was concentrated beyond the passband of many V1 and LM neurons (mean high-cutoffs of 0.07 and 0.055 cycles/degree) (Marshel et al., 2011) and also the differential energy at the orientation domains of the LGs across spatial frequencies was little (Figure 1a,c). Then I compared the orientation tuning curves for LGs and CGs with high-frequency noise carrier across areas V1 and LM. Among all recorded neurons with significant responses to LGs, less than half of them also responded to CGs with high-frequency noise carrier (37%, 51 of 139 recorded neurons; Figure 10). In comparison, CGs with low-frequency noise carrier evoked responses in a larger population (72%, $p < 10^{-11}$, log-linear analysis, interaction). Interestingly, this reduction of responsiveness between two types of CGs stimuli was more prominent in area V1 (77% vs. 36%, 29 of 81 recorded neurons) than in area LM (61% vs. 38%, 22 of 58 recorded neurons; $p = 0.040$, log-linear analysis, interaction; compare Figure 6 and Figure 10).

Again, a similar pattern of results was observed in the responses to LGs and high-frequency noise carrier CGs (Figure 11). In both example neurons and in the population of recorded neurons (Figure 11a,b) responses to CGs with high-frequency noise carrier were lower than to LGs. This reduction was mediated by reduced responses to both the preferred orientation (decrease of 49.2 ± 9.8%, from 13 sp/s to 6.6 sp/s; ANOVA, main effect, $p < 10^{-5}$; Figure 11c,d) and orthogonal orientation (decrease of 29.0 10.4%, from 7.6 sp/s to 5.4 sp/s; ANOVA, main effect, $p = 0.008$; Figure 11e,f). Similarly, d’ dropped

**Figure 9.** Comparison of responses to LGs matched in RMS contrast and CGs. **a**, Responses to LGs with matched RMS contrast (black) and CGs with low-frequency noise (red) of an example neuron from area V1. Top, Raster plot. Bottom, Tuning curves of the same example neuron (Unit 221-1-6.24). **b**, Same, for an LM example neuron (Unit 245-5-7.17). **c, d**, Responses to preferred orientations for CGs versus LGs with matched RMS contrast in the population of (c) V1 neurons, $N = 73$ and (d) LM neurons, $N = 23$. **e, f**, Same as **c, d**, for responses to orthogonal orientations. **g, h**, Same as **c, d**, for $d'$. Conventions as in Figure 7.
considerably by 71.6 ± 6.7% for CGs with high-frequency noise compared with LGs (from 1.34 to 0.38; ANOVA, main effect, p < 10⁻¹³; Figure 11g,h).

Figure 10. Distribution of responsive and unresponsive neurons to CGs with high-frequency noise carrier in area V1 (N = 81) and LM (N = 58).

Figure 11. Responses to CGs with high-frequency noise carrier. a, Responses to LGs (black) and CGs with high-frequency noise (red) of a V1 example neuron. Top, Raster plots. Bottom, Tuning curves of the same example neurons (Unit 326-1-2.48). b, Same for an LM example neuron (Unit 299-3-13.42). c, d, Responses to preferred orientations for CGs with high- frequency noise versus LGs in the population of V1 neurons, N = 29 (c) and LM neurons, N = 22, (d). e, f, Same as c, d, for responses to orthogonal orientations. g, h, Same as c, d, for d'. For other conventions, see Figure 7.

Preferred orientations for luminance-modulated and contrast-modulated gratings are broadly similar

Finally, to test the potential contribution of CG representation in the cue-invariant perception of stimulus orientation, I compared the neurons’ preferred orientation, separately for each grating type (Figure 12). Given the fact that stronger orientation selectivity will entail a more accurate estimate of preferred orientation, I considered only those neurons with d’ > 1 for LGs (Figure 12, black circles). First, I examined the degree of uniformity in the distribution of differences in preferred orientations for LGs and CGs. I observed that the distribution of differences in preferred orientation between LGs and CGs with low-frequency noise (Figure 12 a,b) was not uniform in both area V1 (Rayleigh test, p < 10⁻¹⁰) and LM (Rayleigh test, p = 0.03). Indeed, in both areas V1 (0.43, p < 10⁻⁴) and LM (0.4, p = 0.02), preferred orientation for CGs with low-frequency
noise and LGs were correlated. However, the residual activation of luminance-sensitive mechanisms by the CGs with low-frequency noise also could give rise to similar results. Interestingly, for CGs with high-frequency noise (Figure 12c,d), the differences in preferred orientation were distributed non-uniformly only for area LM (Rayleigh test, $p = 0.005$) and correlations between preferred orientations were observed only for area LM (0.58, $p = 0.02$). This finding is remarkable, as LM neurons, with their preferences for lower spatial frequencies (Marshel et al., 2011), are expected to be less sensitive to any residual orientation signal potentially present at higher spatial frequencies. As the last step, I performed control experiments to assess how much of the observed scatter of preferred orientations arises from estimation errors due to limited data. In these control experiments, only LGs were presented and all other aspects of experiments and analyses were left identical (Figure 12e,f). I observed that only little variability arose from limited data, at least for the strongly tuned neurons. Together, my finding of broadly similar preferred orientation between grating types provides some evidence for a coarse cue-invariant, which might partially be the neural correlates for perceptual generalization of orientation discrimination.

**Figure 12.** Comparison of preferred orientation for contrast-modulated versus LGs. a, Top, Preferred orientations of V1 neurons in response to CGs with low-frequency noise carrier versus LGs. Bottom, Distribution of differences in preferred orientations. Black represents neurons with $d' > 1$ for LGs ($N = 110$). Gray represents neurons with $d' < 1$ ($N = 68$). Square markers represent neurons with $d' > 1$ for both grating types. b, Same as a, for LM neurons ($N = 32$ with $d' > 1$ and $N = 37$ with $d' < 1$). c, Same as a, but for CGs with high-frequency noise carrier versus LGs ($N = 110$). d, Same as c, for LM neurons ($N = 16$ with $d' > 1$ and $N = 13$ for $d' < 1$). e, Same as a, but repeating the LG condition ($N = 53$ with $d' > 1$ and $N = 70$ with $d' < 1$). f, Same as e, for LM neurons ($N = 11$ with $d' > 1$ and $N = 7$ with $d' < 1$).

**Responses to contrast-modulated gratings during anesthesia**

I also performed extracellular recordings under anesthesia, in which I measured the orientation tuning in response to LGs and CGs with low-frequency noise carrier across both areas V1 and LM. The pattern of results was similar to those during wakefulness, such that both V1 and LM neurons were less responsive and selective to CGs than LGs. Indeed, about half of all recorded neurons in both V1 and area LM did not respond to CGs despite their significant responses to LGs (Figure 13). The fraction of responsive neurons to both LGs and CGs (49%) was lower compared to during wakefulness (72% $p < 0.001$, interaction, log-linear analysis). The reduction of responsiveness with brain state was similar for LM (awake: 61% vs. anesthetized: 42%) and V1 (awake: 77% vs. 68%).
anesthetized: 64%; three-way interaction, p = 0.85). The fraction of responsive neurons to second-order stimuli during anesthesia was higher in V1 (64%) than in LM (42%, p < 0.001, interaction, log-linear analysis; Figure 14).

Among those neurons with significant responses to both LGs and CGs, example neurons in both areas V1 (Figure 15a,c) and area LM (Figure 15b,d) had lower peak firing rates for CGs compared to LGs. This reduction of firing rates to CGs was also evident in the population of recorded neurons, where I used an ANOVA to test statistical significance. Responses across recorded areas dropped by 36.1% ± 5.2, from 9.8 spikes/s in response to LGs to 6.3 spikes/s in response to CGs (ANOVA, main effect, p < 10^{-6}, Figure 15e-h). I also noted that the responses to the orthogonal orientation for CGs versus LGs decreased in both areas (area V1, mean change 25.3% ± 8.3; area LM, mean change 21.1% ± 5.7; main effect, p < 10^{-4}).

This decrease in responsiveness was accompanied by a concomitant drop in selectivity for stimulus orientation for CGs, both in the example cells (Figure 15a-d) and in the population during anesthesia (Figure 15i-l). d’ was lower during anesthesia than wakefulness (main effect, p < 0.01). More importantly, d’ during anesthesia was higher for LM than V1 (main effect, p = 0.03) and lower for CGs than LGs (main effect, p < 10^{-5}), and this difference tended to be stronger for LM (69.7% ± 13.9 from 1.22 for LGs to 0.36 for CGs) than V1 (57.3% ± 12.9) from 0.76 for LGs to 0.32 for CGs, interaction, p = 0.05). Thus, responses to CGs compared to LGs in mouse visual cortex during anesthesia are lower and less selective for orientation.

Figure 13. Example neurons unresponsive to CGs despite significant responses to LGs during anesthesia. a, Activity evoked by LGs (black) and CGs with low-frequency noise (red) of an example neuron from area V1. Top, Raster plots. Each gray horizontal line separates trials with different stimulus orientations. Bottom, Tuning curves of the same example neuron (Unit 232-5-8.23). Dashed horizontal line indicates response to gray screen; solid lines indicate fit of a sum-of-Gaussians model. b, Same, for an LM example neuron (Unit 242-2-12.35).
**Figure 14.** Distribution of responsive and unresponsive neurons to CGs with low-frequency noise carrier in area V1 ($N = 55$) and LM ($N = 100$) during anesthesia.

**Figure 15.** Responses to CGs with low-frequency noise carriers in mouse visual cortex during anesthesia. **a,** Responses to LGs (black) and CGs (red) of an example neuron from area V1. Top, Raster plots. Bottom, Tuning curves of the same example neuron (Unit 232-5-4.6). **b,** Same, for an LM example neuron (Unit 144-2-20.39). **c, d,** Response to the preferred orientation for CGs versus LGs in the population of responsive neurons recorded from V1, $N = 178$ (c) and LM, $N = 69$ (d). **e, f,** Same as c, d, for response to orthogonal orientation. **g, h,** Same as c, d, for $d'$. Conventions as in Figure 7.
Discussion

In this thesis, I investigated whether mice can use second-order stimuli in a cue-invariant way to guide visual perception during an orientation discrimination task, and screened for potential neural correlates in mouse visual cortex. I found that mice could readily generalize orientation discrimination learning from the familiar cue condition of luminance-modulated gratings (LGs) to a novel cue condition of contrast-modulated gratings (CGs), albeit with a substantial decrease in performance. My electrophysiological data supported these behavioral results: in both areas V1 and LM, the population of responsive neurons to CGs was less than that of LGs. Furthermore, responses to CGs were generally weaker and less orientation selective than to LGs. Despite these differences, preferred orientations were broadly similar in response to both LGs and CGs, pointing to a potential underlying mechanism of the basic cue-invariant generalization in mouse orientation discrimination performance. My observation of cue-invariance might provide the basis for object recognition and categorization invariant to changes in object appearance, as recently demonstrated in rats (Zoccolan et al., 2009; Tafazoli et al., 2012; Vermaercke and Op de Beeck, 2012; Alemi-Neissi et al., 2013; Vinken et al., 2014; De Keyser et al., 2015).

Choice of paradigm

From previous behavioral studies on the limits of rodent vision, it has become clear that the particularity of the paradigm can strongly influence the measured visual abilities. For example, Busse and her colleagues in 2011 (Busse et al., 2011) used a nose-poke two-alternative forced task (2AFC) to measure mice contrast sensitivity at preferred spatial frequency and reported 20% contrast threshold, whereas a much lower contrast threshold (2%) was estimated when a yes-no paradigm was used, instead (Histed et al., 2012). The discrepancy present between these two studies might partly be explained by the difference in both the level and the stability of animals’ motivation between 2AFC and yes-no paradigms. In the yes-no paradigm used in Histed’s study not only mice’ performance was close to optimal but more importantly it was stable across different sessions of training, as well. Indeed, measuring psychometric functions in an operant, lever-pressing paradigm provided the opportunity to detect any changes in animals’ motivation, arousal or attention and it revealed negligible, near to zero, lapse rates. Moreover, their estimates of threshold were repeatable across sessions and mice, indicating stability in performance and little deviation (if at all) from optimal behavior. In contrast, in the 2AFC paradigm used by Busse et al., mice’ performance was influenced by non-sensory factors such as past history of failures and rewards and estimates of reward values and thus by fluctuations in animals’ motivation. To sum up, estimates of mouse contrast sensitivity in a nose-poke 2AFC task do not exclusively reflect perceptual limits, as it can also be limited by non-sensory factors (Busse et al., 2011), whereas rigorous psychophysical measurements using a lever-press paradigm under head fixation can reveal perceptual thresholds that are less influenced (Histed et al., 2012).

Why did I choose classical conditioning?

In order to test whether mouse vision is capable of perceiving CGs, despite its low acuity, I chose a classical conditioning paradigm. Following Gallistel et al. (2004), and by
performing a quantitative analysis on conditioned responses (LIs) I was able to compare mouse behavior before vs. during CGs presentations. The positive slope of cumulative records of conditioned responses (LIs) indicated that mice have stronger conditioned responses during CGs presentation than before, which simply confirms CGs perception. Furthermore, classical conditioning allowed me to test separately CGs perception and CGs orientation discrimination abilities in mice. The difference between LIs to rewarded and unrewarded CGs orientations, demonstrated orientation discrimination capability of mice for CGs, albeit with a substantial drop in performance. Therefore, employing classical conditioning provided, for my study, the advantage of a distinct behavioral read-out for the visibility of the stimulus and discriminability of its orientation.

Can classical conditioning paradigm reveal the limits of visual performance?
Despite offering the advantage outlined above, classical conditioning might not be able to measure near-threshold performance in my task, due to potential fluctuations in motivational states of animals. As explained earlier, any fluctuation of animal motivation could strongly influence on experimental results and be misleading to estimate visual limits (Busse et al., 2011; Histed et al., 2012). In the other paradigms such as touch screen panels (Bussey et al., 2001), and operant lever-pressing (Histed et al., 2012) which require mice to initiate the trial or participate actively in the experiment to obtain reward, the animals’ attention and motivation might be more stable, the animals’ performances might be less contaminated and might reflect accurately the perception limits. In contrast, classical conditioning, despite offering powerful insights into several aspects of visual perception, can involve more reflexive behavior and trial-to-trial fluctuations in motivation cannot be excluded. Therefore, I speculate that using other paradigms, such as operant lever-pressing (Histed et al., 2012) or touch screen panels (Bussey et al., 2001), might result in a better performance for orientation discrimination of CGs compared to what I observed during classical conditioning in this project. Indeed, a reasonable behavioral generalization of first- to second-order stimuli was observed recently in rats trained in a touch-screen paradigm (De Keyser et al., 2015). In addition, one type of invariant visual object recognition has been reported during an operant paradigm, where rats had to recognize the objects despite a range of changes in size and positions. In this type of paradigm, rats were trained to initiate each trial by inserting their head into a whole in front of training box (Alemi-Neissi et al., 2013).

It is important to note that despite my finding that mice are able to generalize the learned orientation discrimination task from LGs to CGs, this generalization was rather limited. First, not only did discrimination performance substantially drop from LGs to CGs, but extensive training also did not lead to a qualitatively better performance during CGs. Second, I tested one mouse who was never able to reliably discriminate CGs even after prolonged sessions of training. Third, I realized that the wave shape of LGs determines successful transfer of learning to CGs later. Indeed, mice were able to generalize from LGs to CGs only after being initially trained with sine-wave LGs, but never after being initially trained with square-wave LGs. All 3 mice trained initially with square-wave LGs, could no longer discriminate orientation when the gratings were contrast modulated, despite being able to see the grating, as they showed unspecific
conditioned responses to CGs’ presence. Indeed, these mice could see CGs but failed to discriminate different orientations.

**Underlying mechanism of second-order**

As I showed in this study, neurons in both area V1 and LM not only respond to CGs but also are orientation-selective for CGs. Furthermore, comparing the preferred orientation of V1 and LM neurons to LGs and CGs revealed a coarse correspondence between them. Therefore, my electrophysiological data could potentially underlie the perceptual generalization of orientation discrimination from LGs to CGs. Previous studies in higher-order mammals guided by rich knowledge of the spatiotemporal properties across visual areas (Movshon et al., 1978; Issa et al., 2000) suggested a two-stage filter model for second-order processing. In this filter-rectify-filter model, the first linear spatial filter detects the fine structures of second-order stimuli and then its output is transferred to a nonlinear rectifier, and finally, the second linear spatial filter sums up the rectified responses. However, the neural substrates underlying these three stages remain controversial, particularly those pertaining to the first filter. An electrophysiological study in area V2 of monkeys has proposed that small receptive fields in V1, tuned to spatial frequency and orientation, act as a potential neural substrate for the first linear filter (Li et al., 2014), while another study in cats (Rosenberg et al., 2010; Rosenberg and Issa, 2011) reported nonlinear responses and tuning to carrier spatial frequency and orientation of CGs in LGN Y-cells, suggesting that subcortical structures, i.e. LGN, could also serve as first stage filter. Regarding the neural substrate of the second filter, it has been suggested that V2 neurons which have bigger receptive fields than V1 neurons and consist of V1 receptive fields sum up the rectified responses coming from primary visual area’s RF and provide the orientation selectivity (Li et al., 2014). However, it is uncertain yet whether primate or carnivore brain possesses such second-order processing stages or whether mechanisms other than filter-rectify-filter, such as surround-suppression, explain the second-order responses (Tanaka and Ohzawa, 2009; El-Shamayleh and Movshon, 2011; Hallum and Movshon, 2014; Li et al., 2014). Tanaka and Ohzawa (2009) suggested that neural responses to a second-order stimuli in area 17 of cats arise from surround suppression, however, measuring the suppression index of V2 neurons showed little or no surround suppression for half of second-order responsive neurons. Furthermore, employing high ratio of carrier spatial frequency to envelope frequency indicated that the second-order responses could potentially be explained by filter-rectify-filter mechanism instead (Li et al., 2014). My findings of cue-invariant responses to CGs in mice cannot resolve this debate, as the stimulus I used cannot determine accurately the source of those responses. The observed cue-invariant responses to CGs used in this study could potentially emerge either from residual activation of luminance-sensitive mechanisms or could be mediated by suppressive mechanisms from the surround (Tanaka and Ohzawa, 2009; Hallum and Movshon, 2014). Nevertheless, the pattern of LM responses to high-frequency noise carrier CGs is reminiscent of the filter-rectify-filter framework (see below).

**Potential role of LM in perception of texture boundaries**

The electrophysiological data that I acquired in this project points towards a potentially interesting role of area LM in the perception of texture boundaries. As I showed in this study, LM neurons in mice not only respond to CGs with high-frequency noise carriers,
but also maintain some orientation selectivity to these CGs. More importantly, unlike in V1, neurons in area LM showed a broadly similar preferred orientation for LGs and CGs with high-frequency noise carriers. This finding is remarkable because first, it is not probable that the high spatial frequency distribution of the noise carrier evokes the luminance-sensitive mechanism and second, area LM should be less sensitive to any residual orientation signal potentially present at higher spatial frequencies (Marshel et al., 2011). These characteristics of LM neurons offer the potential involvement of area LM in the perception of texture boundaries. This potential role of area LM is particularly intriguing, as the properties of area LM fit well with the filter-rectify-filter model: similar to V2 and area 18 in monkeys and cats, LM has been indicated to be a primary projection target of V1 (Levitt et al., 1994; Wang et al., 2011; 2012), has bigger receptive fields (Levitt et al., 1994; Price et al., 1994; Van den Bergh et al., 2010) and shows lower spatial frequency tuning compare to V1 (Movshon et al., 1978a; Foster et al., 1985; Marshel et al., 2011). However, to test explicitly whether filter-rectify-filter mechanisms exist in the mouse visual cortex, further experiments are required to assess precisely the carrier spatial frequencies. In order to exclude any source of luminance-sensitive mechanism in cue-invariant neuronal responses, a carrier with high spatial frequencies unresolvable by the first-order mechanism needs to be employed. In such experiments, one should determine tuning for carrier spatial frequency and aim to rule out any possible source of nonlinear artifacts, which might come either from the display device or the photoreceptors (Zhou and Baker, 1994; Mareschal and Baker, 1998a; Li et al., 2014). These various artifacts either cannot yield such carrier tuning or if they can, it would be the same in every neuron. Therefore, tuning for spatial frequency, which varies across neurons and is distinct from that of the envelope, could ensure that the observed nonlinearity does not emerge from artifacts. In addition, being able to demonstrate a spatial frequency tuning to the carrier distinct from that predicted by a surround mechanism (Tanaka and Ohzawa, 2009) would rule out activation of luminance-sensitive mechanisms (Li et al., 2014).

Are the other mouse extrastriate areas more strongly responsive to second-order stimuli?

I demonstrated that area LM is responsive and orientation selective to second-order stimuli, however, it is not clear whether any of the other mouse extrastriate areas respond more strongly to second-order stimuli. Currently, it is debated whether area LM belongs to the ventral or dorsal stream, despite being identified as a V2 homologous due to its distinct connections and the shared representation of the vertical meridian (Coogan and Burkhalter, 1993; Wang and Burkhalter, 2007). Studies based on cytoarchitectonic and chemoarchitectonic markers and pathway tracing suggested that area LM belongs to ventral stream as it receives its strongest projections from V1 and forwards inputs to limbic areas (e.g. the lateral entorhinal cortex and the amygdala) and also to the superficial layers of the superior colliculus (Wang et al., 2011; Wang and Burkhalter, 2013). Despite being a powerful approach for exploring the interconnectivity and potential function of visual processing streams (Wang and Burkhalter, 2007; Wang et al., 2012), pathway-tracing methods hold some limitations. In fact, the strength of projections to a certain area may not represent precisely the function of neurons in that area. This is because it is not just the number, but also the strength and functional contents of synapses which determine the physiological role of a given structure (De Pasquale and Sherman, 2011). On the other hand, measuring
fundamental properties of extrastriate areas in mouse visual cortex suggests that detailed structure of visual images are analyzed, not by LM, but rather other areas preferring high spatial frequencies such as PM and LI (Andermann et al., 2011; Marshel et al., 2011). Using two-photons calcium imaging, Marshel et al. (2011), measured the fundamental properties of mouse visual areas and suggested that area PM with its high spatial frequency preferences and high orientation selectivity could resemble the primate ventral stream. They also proposed a similar but less dominant role for LI area. However, their results in area LM were less compelling, as LM showed the highest temporal frequency (1.8 HZ) and a moderate spatial frequency (0.028 cycles/ degree) among seven visual areas, which brought them to conclude that area LM is more dorsal-like than ventral-like area. In addition, another study comparing responses to pattern and component motion across different visual areas showed that area LM responds to pattern motion, reminiscent of the dorsal stream, which is in contrast to V1 and AL areas, which are more component-like (Juavinett and Callaway, 2015). To sum up, the functional division of rodent extrastriate areas is a matter of debate and despite the established similarities of mouse vision and primates, it is difficult to assign distinct tasks to each individual area in mouse visual cortex. A recent study in rats investigated the correlation of neural responses in extrastriate areas and behavioral performances in a shape discrimination task and reported that neural responses in higher visual areas represent more behavioral discriminability, while neural responses in V1 area were more closely related to physical differences between visual stimuli (Vermaercke et al., 2015). It is, therefore, likely that other extrastriate areas, which are located in higher levels of the mouse ventral stream hierarchy, respond more strongly to second-order stimuli, as in nonhuman primates (Poort et al., 2012; An et al., 2014). In agreement with this notion, studies in primates which examined motion perception (An et al., 2012) and figure-ground segregation (Poort et al., 2012) across V1 and higher visual areas observed more prominent and stronger responses in V2 and V4 areas than V1 area.

**It is also currently unknown whether responses to second-order gratings are stronger during task performance.**

Whether performing a task could enhance the second-order responses is not clear. The ideal experiments would be those in which neural responses to both grating types were measured during the orientation discrimination task in both naive and trained animals. A recent study showed that learning of visual discrimination with first-order gratings modulates the population responses in area V1 (Poort et al., 2015). Indeed, by imaging the population neurons in layer 2/3 of V1, they found that learning enhances the representation of task-relevant stimuli in the population-level in consecutive training sessions. Tracking individual neurons during consecutive training sessions demonstrated that the observed enhancement of neural representations at the population-level occurred via stabilizing the existing and recruiting new neurons selective for task-relevant stimuli. In fact, learning decreases the variability of a given neuron in its selectivity for a certain stimulus (Peters et al., 2014) and in addition, recruits newly selective neurons, which were not selective before learning (Poort et al., 2015). It is also possible that similar alterations occur during learning of other grating types such as second-order stimuli.
Conclusion
To summarize, in this thesis I have examined the processing of different second-order visual stimuli in the mouse model. I found that mice are capable of generalizing orientation discrimination learning from first-order to second-order gratings and their neurons in both V1 and area LM are orientation-tuned to second-order gratings. These results are remarkable, as they offer new insights for future studies to investigate simple forms of invariance and their circuit-level neural mechanisms in the mouse model, where various genetic engineering methods are available for chronic imaging of all visual cortical areas (Adermann et al., 2011; Marshel et al., 2011), causal manipulation of specific cell types (Fenno et al., 2011), and circuit tracing (Wickersham et al., 2007).
Acknowledgment

Firstly, I would like to express my special thanks to my advisor Prof. Laura Busse for the continuous support of my PhD research and for her patience, encouragement and immense knowledge. She has been always academically and emotionally supportive and helpful in last 5 years. Beside my advisor, I would like to thank my committee members, Prof. Thomas Euler, Prof. Cornelius Schwartz and Prof. Uwe Ilg for serving as my committee members even at hardship. A special thanks goes to Prof. Horst Herbert, Prof. Matthias Bethge and Prof. Andreas Bartels for their supports during my PhD. I would like to thank Dr. Reza Daliri, who provided me the opportunity to join Prof. Laura Busse’s lab. Without his support would not be possible to conduct this research. I am also grateful to Dr. Steffen Katzner, lab-mates and friends for their kind supports. Last but by no means least, I thank my great family who has been always supportive in my life and during my PhD as well.
References


