



# Similarity analysis of brightness perception effects for multiscale spatial filtering models

by

**Hendrik Schulze Bröring**

**Matriculation Number 401889**

A Bachelor's thesis submitted to

Technische Universität Berlin  
School IV - Electrical Engineering and Computer Science  
Institute of Computer Engineering and Microelectronics  
Computational Psychology

November 27, 2022

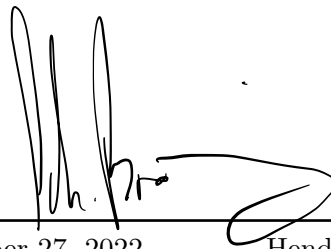
Supervised by:  
Prof. Dr. Marianne Maertens

Assistant supervisor:  
Prof. Dr. Guillermo Gallego

## Eidesstattliche Erklärung / Statutory Declaration

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

I hereby declare that I have created this work completely on my own and used no other sources or tools than the ones listed.



---

Berlin, November 27, 2022

Hendrik Schulze Bröring

# Acknowledgments

I would like to thank Dr. Joris Vincent for his guidance throughout the whole project.

# Abstract

Multiscale spatial filtering models predict human brightness perception of an image; oftentimes of a brightness perception effect (stimulus). The application of the models are computationally expensive and slow. To increase their performance, one can decrease the pixel quantity of the stimulus it is applied to. However, a decrease in pixel quantity (resolution) may change properties of the brightness prediction. To quantify that change, image properties of the default brightness prediction are compared to the same properties of predictions with different resolutions. To avoid idiosyncrasy, the brightness prediction of seven types of stimuli and multiple resolutions are assessed. In fact, all the brightness prediction properties that were analyzed, deviated from the default prediction. Even though, some properties remained more stable than others, the exact change was dependent on the stimulus to which the model was applied.

# Zusammenfassung

Multiscale spatial filter Modelle prognostizieren die menschliche Helligkeitswahrnehmung eines Bildes; häufig handelt es sich bei dem Bild um einen Helligkeitseffekt (Stimulus). Die Anwendung der Modelle ist rechenintensiv und langsam. Um ihre Leistung zu steigern, kann man die Pixelanzahl des Stimulus, auf den sie angewendet werden, verringern. Eine Verringerung der Pixelanzahl (Auflösung) kann jedoch die Eigenschaften der Helligkeitsvorhersage verändern. Um diese Veränderung zu quantifizieren, werden die Bildeigenschaften der Standard-Helligkeitsvorhersage mit den gleichen Eigenschaften von Vorhersagen mit unterschiedlichen Auflösungen verglichen. Um Idiosynkrasie zu vermeiden, wird die Helligkeitsvorhersage von sieben Arten von Stimuli mit jeweils mehreren Auflösungen bewertet. Es wurde gezeigt, dass sich mit veränderter Auflösung auch die Eigenschaften der Helligkeitsvorhersagen veränderten. Obwohl einige Eigenschaften stabiler sind als andere, hängt die genaue Veränderung von dem Stimulus ab, auf den das Modell angewendet wurde.

# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Methods</b>	<b>11</b>
2.1	Stimuli . . . . .	12
2.2	Resolution and size . . . . .	13
2.3	Spatial filtering model and model responses . . . . .	14
2.4	Comparing model response images . . . . .	16
2.5	Resources and tools . . . . .	17
<b>3</b>	<b>Results</b>	<b>18</b>
3.1	Distribution of model response intensity values . . . . .	18
3.2	Power spectrum analysis . . . . .	20
<b>4</b>	<b>Discussion</b>	<b>25</b>

# 1

## Introduction

The brightness of surfaces in the visual environment is one of the properties detected by the human visual perception apparatus [1]. It is well-known that the human brightness perception (psychological intensity) is not only determined by the luminance (physical intensity) of a surface, i.e., the light waves it reflects[2], [3]. In fact, an additional factor that influences the human perception of brightness is the luminance of the surface's surrounding. That circumstance can lead to visual effects in which, e.g., two patches with equal luminance are perceived with different brightness, depending on the surrounding context. Such a brightness effect can be observed for Whites' Effect[4], which shows two gray patches that are identical in luminance, but are perceived to have different brightness in different surrounding contexts(Fig. 1.1a). Some of these brightness effects are attributed to properties of retinal ganglion cells of the human visual system: Simplified, these cells consist of an inner and an outer circle. The cell is activated when light hits the inner circle and inhibited when light hits the outer circle[5].



(a) Brightness perception effect, named White's-Effect[4]. The luminance of the gray patches is identical. The patch that shares more edges with the white stripes appears brighter.

(b) Response of the FLODOG model that predicts the human brightness perception for White's-Effect.

Figure 1.1

Spatial filtering models that aim to explain human brightness perception [1], [3], [6] have been the subject of academic investigations for decades. These models can be applied to achromatic digital images of brightness perception effects (stimuli), such as White's Effect, and respond with a prediction of the human brightness perception of that stimulus (Fig. 3.4b). The models use Gaussian filters, which mimic the aforementioned cell behavior [7]. The images to which these filters are applied, are composed of a limited number of elements, also referred to as pixels [8, p.18]. Each pixel has a location

and a numerical intensity value (also gray level value) between 0(black) and 255(white)[8], [9]. In that sense, a digital image is represented as a matrix and each pixel represents a matrix element. The filters are represented as matrices as well (Fig. 1.2). During the execution of a model, a filter is applied to each pixel of the image and yields a new value by performing a mathematical operation on the pixel value and its respective neighboring pixels [p.154][8].

A spatial filtering model consists of multiple filters. The FLODOG model, which has performed remarkably well in predicting perceived brightness of an image, consists of 42 Gaussian filters[1]. The many filters and respective computations cause the model to be very time-consuming. The larger an image, the larger the matrix, the higher the pixel quantity, the more computations have to be performed by the filter, the more time-consuming is the execution of the model. However, the execution time can be significantly reduced, by reducing the size, i.e. the amount of pixels, of the image it is applied to. If the size in pixels of White's-Effect (Fig. 1.1a) is reduced by factor 4, the respective model execution time is reduced from 31 seconds to 6.48 seconds (Fig.1.3). For users of the model who apply it to several stimuli, multiple times a day, the time saved can be significant.

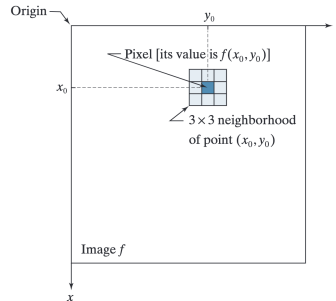


Figure 1.2: Spatial filter that operates on the pixel and its corresponding 3x3 neighborhood. Source: Gonzalez and Woods [8, p.120]

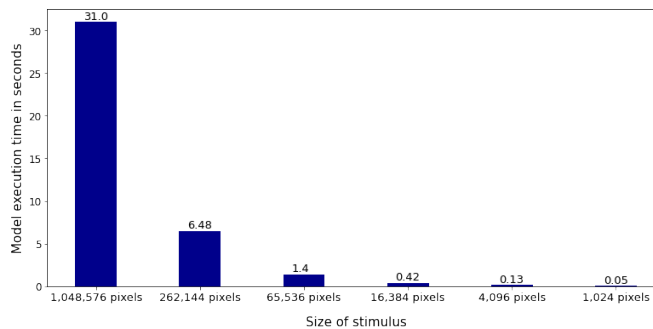


Figure 1.3: Execution time of the FLODOG model[1] applied to White's-Effect. The model was applied to varying sizes of the stimulus.

Reducing the size of the stimulus might cause a model response that differs significantly from the responses of the original stimulus. However, it is not very straightforward to predict which differences occur at which pixel sizes. From a purely visual perspective, model responses with different pixel sizes can appear very similar and very different (Fig. 1.4). Even though, some model responses do not differ visually, they must differ in some way if their pixel size differs: If the amount of pixels is reduced, loss of information is inevitable. Especially, when the change of model responses is invisible, it is crucial to understand how the loss of information changes the model response. Anyone who works with these models might rely on visual and non-visual properties of the model response to be stable. Moreover, due to the many filtering and normalization steps, it is non-trivial to predict these changes in model response before the model execution. Hence, if one wants to leverage smaller stimuli to speed up the model execution time, it is critical to answer the question of how this reduction affects the model responses.



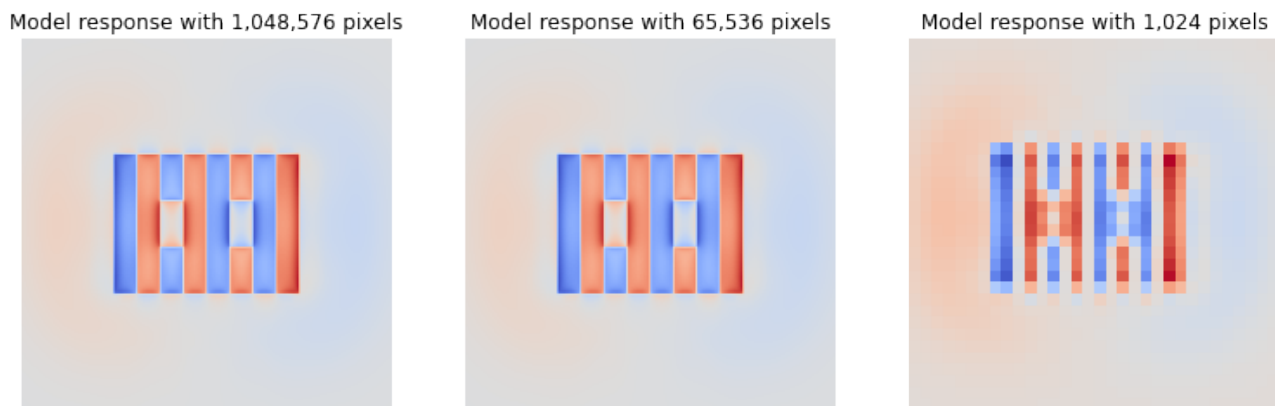


Figure 1.4: Response of the FLODOG model applied to White's-Effect with 1,048,576, 65,536 and 1,024 pixels. The model response on the left and in the middle appear very similar despite their huge difference in pixels. Despite their visual appearance, they must be different in some way. The left image contains more pixels, and hence, more information than the image in the middle. The right model response obviously differs from the left and the middle one.

To analyze how reduced model responses differ from the original size, several brightness perception effects (stimuli) with respective size-differing variations are created. The spatial filtering model will be applied to the original size stimulus and the variations that differ in size. The original model response will then be compared to the other model responses to evaluate the difference. A difficulty for the evaluation is that the model responses consist of different pixel sizes. If they were the same pixel size, the comparison would be less difficult. One could just subtract one model response from the other. Well-established metrics that allow for comparison of digital images, e.g. the model responses, exist for decades: For example, analyzing the distribution of intensity values is a well-established method to compare image differences and similarities[10]–[12]. The distribution of intensity values quantifies the amount of pixels per intensity value, i.e. for each intensity value, it indicates the occurrence of that value in the corresponding image of the model response. Another widely utilized method for digital image comparison is the power spectrum[13], [14]. The power spectrum is a common way to analyze the frequencies contained in an image and their amplitudes [15], i.e. it captures the magnitude and rate of alternations between intensity values of the model response. Both methods, the intensity value distribution and the power spectrum, are not only exceptionally well established, they complement each other very well: The distribution is suitable for the analysis of non-spatial information. It captures the quantity of different intensity values, irrespective of their location. The power spectrum, however, assesses the alternations between pixels, i.e. it covers the spatial relationship (spatial information) between intensity values. In conclusion, the distribution of intensity values and the power spectrum are utilized to perform a comparative model response image analysis, from which one can derive the implications of decreasing the stimulus pixel size to reduce the execution time of spatial filtering models.

Indeed, the results of the analysis show that decreasing the size of the stimuli changes properties of the corresponding model responses. However, the degree of change is very dependent on the difference on the individual stimulus and its size in pixels. The similarity analysis produced four key findings: With decreasing size, the distribution smooths out, i.e. the distribution variability decreases. Furthermore, the range of intensity values varies in a non-predictable manner, i.e., the minimum and maximum intensity values differ significantly when the model response pixel size changes. The average intensity value remains close to zero, irrespective of the model response pixel size. However, larger model responses converges to 0, i.e., the distribution of intensity values is centered around zero. More-

over, the power spectrum analysis reveals that higher spatial frequencies are lost when pixel size of the model response is decreased. But not only high frequency information lost, model responses differ in low frequencies as well. In conclusion, if one wants to utilize smaller stimuli to increase the model execution time, one has to be aware of a variety of model response properties that change.

## 2

# Methods

Applying a spatial filtering model is very time-consuming. This is due to the many filter operations that are performed on every pixel of the model input image (usually a stimulus). However, if one reduces the amount of pixels of that stimulus, the execution time of the spatial filtering is drastically reduced. For users of the model, the time saved can be significant. However, a reduction causes changes in the model response. Predicting or observing these changes is non-trivial. Yet, it is crucial to be aware of model response changes, as anyone who works with them relies on some of their properties to remain stable. To evaluate the (dis)similarity of model responses, one can utilize the distribution of intensity values and the power spectrum. Both methods are well-established in image comparison and complement each other very well. In conclusion, the goal of this thesis is to evaluate the change of model responses. The methodological approach to the goal is to take a stimulus; create the stimulus in various pixel sizes; apply the model to each variation to get the model response; and to compare the model responses (Fig. 2.1). This chapter elaborates on the definitions, representations and technical implementation of the stimuli, its size and resolution, the spatial filtering model, the model responses and the measures to compare them.

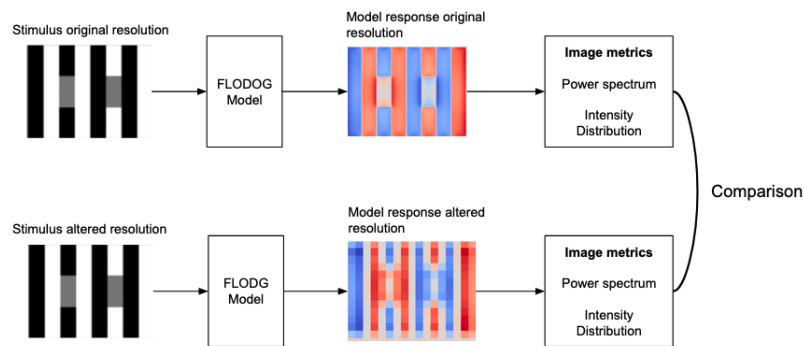


Figure 2.1: The stimulus, e.g. White's Effect, serves as an input to a spatial filtering model, e.g. the FLODOG model. The model output is a digital image of the predicted brightness perception of the stimulus. The model is applied to a stimulus of original pixel size, and a stimulus of altered pixel size. To analyze the (dis)similarity of the respective model responses, the power spectrum and intensity value distribution of the model responses are used to compare them.

## 2.1 Stimuli

Seven stimuli are considered for the comparative analysis (Fig. 2.2). It is necessary to use various stimuli to reduce the likelihood that analysis results are not idiosyncrasies of a single stimulus. The selection of stimuli includes two types of brightness perception phenomena: Assimilation and contrast phenomena. Assimilation includes “examples where the brightness of a test patch is shifted toward the brightness of the region that it shares the majority of its border with”, the latter characterizes “examples where the brightness of the test patch is shifted away from the region that it shares the majority of its border with”(Robinson, Hammon, and Sa [1]). The stimuli are created following the implementation of Robinson, Hammon, and Sa [1]. The attentive observer will have noticed, that each stimulus is displayed on a gray, rectangular background. Technically, the background and stimulus are different objects. However, Robinson, Hammon, and Sa [1] states that the background is necessary “to allow for valid filtering”. In this thesis, I will refer to the term stimulus to describe the combination of stimulus and background, unless stated differently. For the sake of completeness, it is worth mentioning that the intensity values of each stimulus are normalized. Instead of a gray scale between 0 and 255 as described in ??, they are composed of values between 0 (black) and 1 (white). In chapter 1 the size of a stimulus was introduced as the quantity of pixels it contains. The reality is a little more complicated. The relation between pixel quantity and stimulus size will be explained in detail in section 2.2. In fact, the size of each stimulus is  $32^{\circ} \times 32^{\circ}$  degrees of visual angle[1]. The visual angle describes the relation between object size and its distance to the spectator. In scientific experiments, the stimuli are presented to subjects[1], [3], [6]. To ensure comparability, the stimuli are presented so that they always take up the same amount of space of the subjects’ visual field. Hence, the size of a stimulus is expressed in degrees of visual angle. The stimuli are implemented according to Robinson, Hammon, and Sa [1]. For the original stimuli sources, refer to table ??. The naming of the stimuli is in accordance with Robinson, Hammon, and Sa [1], except for Checkerboard-0.94 and Checkerboard-2.1. Throughout the thesis they will be referred to as Checkerboard-0.938 and Checkerboard-2.09 to reflect their exact implementation details (compare Robinson, Hammon, and Sa [1]).

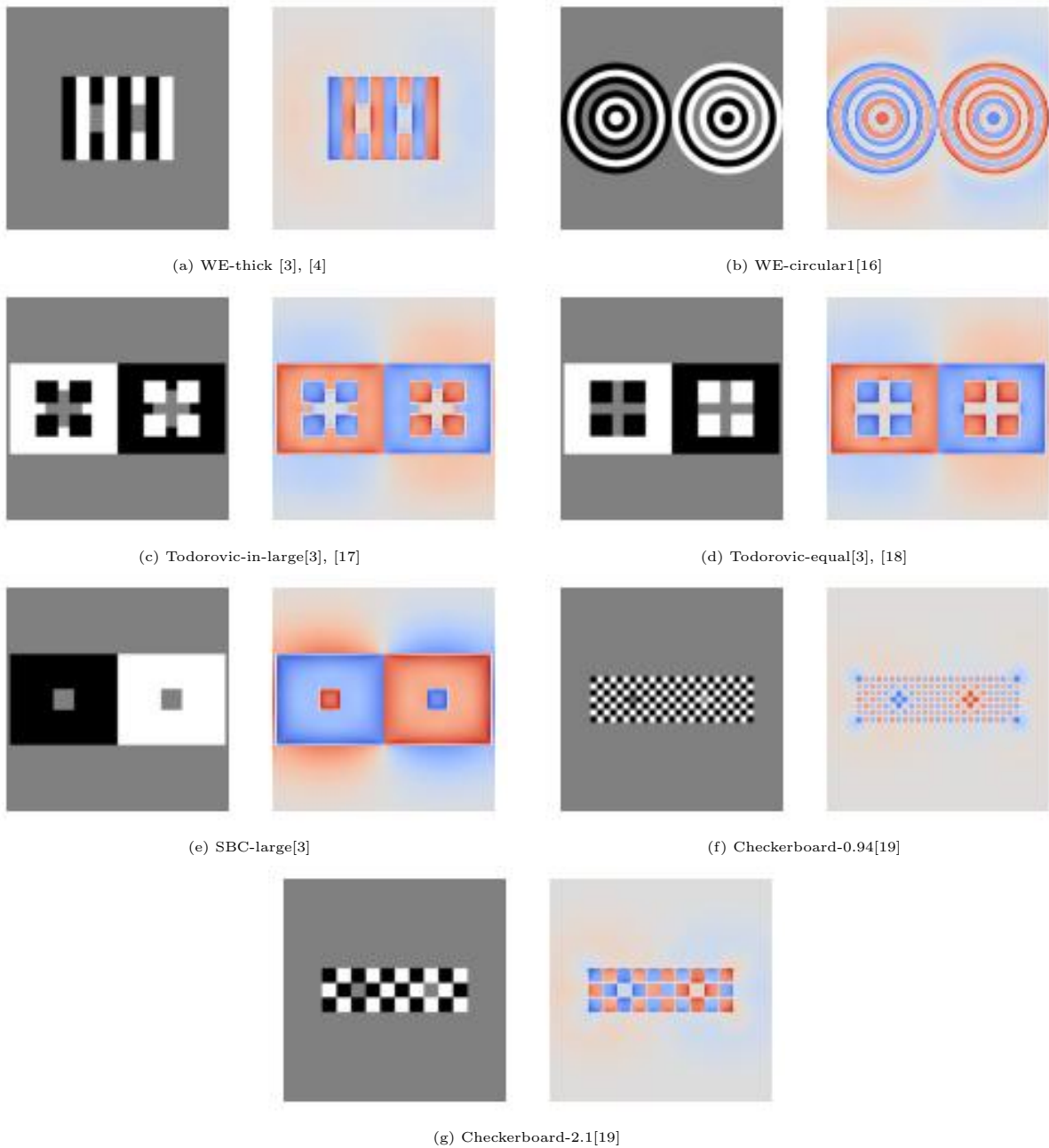


Figure 2.2: This figure displays each stimulus and the corresponding "FLODOG" model response. The model responses are the object of the similarity analysis.

## 2.2 Resolution and size

In chapter 1 and the previous sections of this chapter, I refer to the term size as the pixel quantity of an image. Furthermore, the question this thesis aims to answer was phrased as follows: How the model response changes if the size of the underlying stimulus changes, i.e. how does the model response change with changing quantity of pixels? In section 2.1, the size of the stimulus was re-defined to be

represented in degrees of visual angle. Hence, the term size will not be used anymore to refer to the pixel quantity. This is why it is necessary to introduce the concept of resolution. The spatial resolution of an image is commonly represented as “dots (pixels) per unit distance” [8]. For the stimuli and model responses, the unit distance is degree of visual angle. Consequently, their resolution is expressed as pixels per degree of visual angle (PPD). The higher the amount of pixels per degree, the higher the resolution and vice versa. So, if the pixel quantity of a stimulus changes, it actually reflects a change in resolution, not in size. Hence, correctly phrased, this thesis aims to answer how the model response changes when the resolution of the associated stimulus changes.

The stimuli are used by Robinson, Hammon, and Sa [1] with a resolution of 32PPD. Hence, the “stimuli” package creates stimuli with a default resolution of 32PPD as well. The default model response has the same resolution. To answer how this default 32PPD model response changes with changing resolution, it is compared against model response of the same stimulus that differ in resolution. Because the motivation is to reduce the pixel quantity and consequently the model execution time, only stimuli with lower resolutions are candidates to replace the default resolution and simultaneously speed up the model execution time. However, to understand general patterns, it is helpful to examine higher resolutions as well. This is why, model responses with resolutions of 64PPD, 16PPD, 8PPD, 4PPD, 2PPD and 1PPD are used to perform a comparative analysis against the default 32PPD model response.

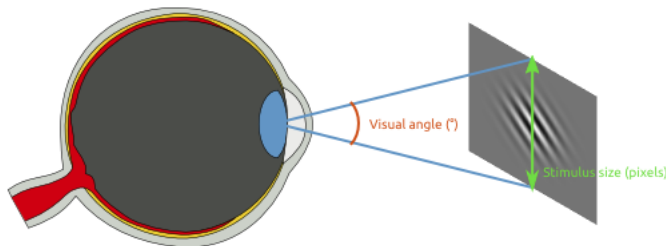


Figure 2.3: The visual angle describes the size of an object, e.g. a stimulus, in relation to its spectator. Source: Mathôt [20]

## 2.3 Spatial filtering model and model responses

As mentioned in section ??, spatial filtering is at the core of several brightness perception models[1], [3], [6], [19]. The essence of spatial filtering is the alteration of intensity values. Gonzalez and Woods [8] describe that in detail: “The spatial domain [alteration] processes [...] are based on the expression  $g(x, y) = T[f(x, y)]$ , where  $f(x, y)$  is an input image,  $g(x, y)$  is the output image, and  $T$  is an operator on  $f$  defined over a neighborhood of point  $(x, y)$ .” The operator is then applied to the pixels of that image. Fig.1.2 illustrates a point  $(x_0, y_0)$  and the corresponding neighborhood. A particular group of spatial filters, the Difference-of-Gaussian(DoG) filters, are used to model components of the human visual apparatus, including brightness perception[1], [3], [6]. The DoG filters are designed to model the behavior of retinal ganglion cells. In simple terms, these cells consist of two concentric circles. If light falls into the inner circle, the cell is activated; if light falls into the outer circle, the cell’s activity is suppressed [7], [21]. The DoG-filter models that “center-surround”(McMahon, Packer, and Dacey [7]) effect. The filter is the sum of a positive, center Gaussian function and a negative, surround Gaussian function[21], [22](Fig. 2.4a). Fig. 2.2e illustrates that the perceived intensity (brightness) of a gray

patch depends on the respective background (either black or white). That suggests, the human visual system combines the intensity value at the gray patch (center) and at the background (surround) to compute brightness. And so does the DoG-filter. One model, that has proven to capture many properties inherent to a variety of brightness perception effects, is the oriented difference-of-Gaussian (ODOG) model by Blakeslee and McCourt [3]. Robinson, Hammon, and Sa [1] elucidate the model’s functionality and processes: It consists of 42 oriented DoG filters that differ in orientation and scale. Oriented DoG filters are much like DoG filters except that they are not circularly symmetric, i.e., either the surround or the center is elliptic[22] (Fig. 2.4b). The outputs of each filter operation are summed, normalized and recomposed to produce the model response which contains a prediction of the stimulus’ perceived brightness[1]. However, Robinson, Hammon, and Sa [1] demonstrate that the ODOG-model “fails on [...] variation[s] of White’s effect”. To compensate for that weakness Robinson, Hammon, and Sa [1] developed a model called FLODOG that introduced a new normalization procedure. This model is used in this thesis. Whenever I refer to the model response, I refer to the FLODOG model response.

The FLODOG model is used to predict the perceived intensity (brightness) of a stimulus. The model takes a stimulus as an input and modifies the corresponding intensity values. The model output is the perceived brightness of that stimulus - the model response. The stimulus is represented as a matrix with an intensity (gray scale) value at each matrix element. Likewise, the model response is a matrix of the same size. Each matrix element of the model response represents the perceived brightness of the stimulus’ intensity value at that same position. The model response is a digital image and consists of intensity values as well. The scale of intensity values range from negative (darker) to positive (brighter) values, because the filters are designed to perform subtractions on positive values and, hence, may result in negative and positive values. Finally, it became apparent how many computations are involved in the application of a multiscale spatial filtering model. That explains why the execution of the model is so time-consuming and why it is so valuable to decrease the stimulus’ resolution and consequently the model execution time.

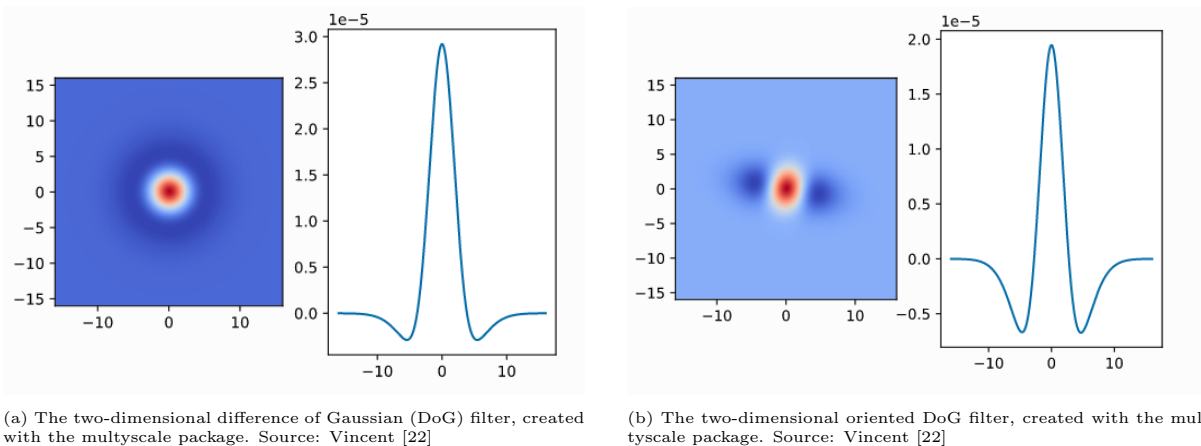


Figure 2.4

## 2.4 Comparing model response images

Decreasing the resolution of a stimulus, to decrease the execution time of the model, introduces potential changes in model response properties. To understand in which ways model responses are (dis)similar from another, the default model response is compared with model responses of various resolutions. The model responses happen to be digital images. Comparing digital images and evaluating their (dis)similarity is a well-established field of research [10]–[13], [23].

Despite similarity analysis being an established field of research, many image similarity measures are developed for images of natural scenes, which the model responses are clearly not. Consequently, one has to be careful when choosing an image similarity measure. Demidenko [9] states that images can be divided into two groups: Structured and unstructured ones. Structured images are easily recognizable by humans, e.g. a “human face, building, carrot, rug, and so on”. Unstructured images are not that easily interpretable by humans (e.g. remote-sensing, medical or astronomy imagery). I argue that model responses belong to the category of unstructured images. As with medical images, humans are not capable of extracting semantics from these images, unless they’re professionally trained. Think of the years of training that a radiologist needs to interpret MRI images. Demidenko [9] argues that unstructured images “may be well represented by its gray level [intensity value] distribution”. The literature provides many examples in which the gray scale (intensity value) distributions are used to compute image similarity[10]–[12].

A distribution of intensity values consists of a horizontal axis which displays the image intensity values, and a vertical axis that displays their absolute and/or relative occurrence (Fig. 2.5). The distribution quantifies the amount of pixels that have a distinct intensity value, but it abandons information to the pixel location[24]. Hence, spatial relationships are ignored. Instead, the distribution contains general information about the scale, distribution, density, and mean of the images’ intensity values[14].

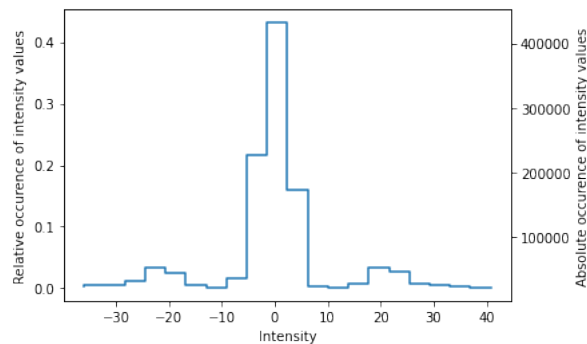


Figure 2.5: Distribution of intensity values for the WE-thick, displaying absolute and relative occurrence of intensity values present in the image

The power spectrum is another feature that is widely used to compute the similarity between unstructured images (see Arun and Menon [13], Wang, Reese, Zhang, *et al.* [23]). To retrieve its power spectrum, an image must be converted from the spatial representation to the frequency based representation. The spatial representation of an image consists of a set of intensity values at specific spatial locations, in other words, a  $M \times N$  matrix where each matrix element represents a pixel. To describe the content of that image as a frequency-based representation, it requires the same  $M \times N$  matrix. However, the values of the matrix elements mean something entirely different: Instead of intensity values and their spatial location, they represent amplitudes and phases of sinusoidal constituents[15,



p.29-31]. A method, which is used to analyze the frequency based representation of an image, is the Discrete Fourier Analysis. Its result is the representation of the image as a relation of frequency and amplitude. The power spectrum is the squared amplitude spectrum. It displays which frequencies exist with what magnitude in a given image[15, p.33-34] and encodes information from basic elements, e.g. orientations, pattern of periodicity and prevalent frequencies[25, p.101].

## 2.5 Resources and tools

Throughout this thesis, Python 3.10.1 is used. The creation of the stimuli is facilitated by the "stimuli" package of the Department of Computational Psychology at Technische Universität Berlin [26]. To apply spatial filtering models to stimuli, I use the "multyscale" Python package, which is provided by the Department of Computational Psychology at Technische Universität Berlin as well[27]. The Python package NumPy is used because of its ability to handle two-dimensional arrays (matrices) and its variety of mathematical functions that allow for extraction of many helpful matrix properties. Matplotlib is a comprehensive Python library that provides many visualizations to analyze and visualize the data. Many images within this thesis were created using Matplotlib. Jupyter notebooks were used to explore the data, assess different approaches, execute computations and plot the visualizations.

# 3

## Results

To analyze how decreasing the stimulus resolution, changes the associated model response, seven types of stimuli are utilized (2.2). For each type, various resolutions are created. In a next step, the FLODOG model is applied to each resolution. The yielded model responses are digital images that predict the perceived intensity of the underlying stimulus. These model responses are compared using two image metric: The distribution of the model response intensity values and the power spectrum of the model response (2.1).

The results of the comparative analysis show that the model responses of with lower resolutions vary from the default model responses. Four key findings describe the variation:

1. The distribution of intensity values smooths out when the resolution decreases.
2. The range of intensity values change depending on the underlying stimulus.
3. The average intensity value converges to 0 with increasing model response resolution.
4. Finding of power spectrum analysis.

### 3.1 Distribution of model response intensity values

The distributions provide a comprehensive overview of how model responses with altering resolutions differ from each other (Fig. 3.1 and 3.2). Three patterns particularly stand out: First, the distribution is relatively symmetric and centered around zero; irrespective of the resolution or stimulus. Second, with decreasing resolution, the minimum intensity value increases and the maximum decreases. Third, the distribution of intensity values smooths out when the resolution decreases.

The range (difference between minimum and maximum) of intensity values seems to decrease systematically with decreasing resolution (Fig. 3.2a, Fig. 3.2b and Fig. 3.2c). Fig. 3.2d suggests that this pattern also occurs for the model response of SBC-large. The model response distributions of WE-thick, Checkerboard-0.938, Checkerboard-2.09 (compare Fig.3.1, 3.2e and 3.2f do not indicate whether the phenomenon also occurs there. A more detailed analysis shows that the relationship between range and resolution is, in fact, non-linear(Fig. 3.3. Despite the general trend that a lower resolution corresponds to a lower range, the exact rate of change is very dependent on the underlying stimulus. Yet, a common pattern among all stimuli emerges: The range peaks at 8PPD or 16PPD. From there, the range decreases again with increasing model response resolution. Indeed, the range does not grow

linearly with increasing resolution. The exact magnitude of decrease and increase between varying resolutions changes between the stimuli, the model was applied to. With a value above 100, the range of Checkerboard-0.938 peaks at 16PPD. In comparison, the model response of SBC-large displays a peak between a value of 60 and 70 at a resolution 8PPD. Moreover, the model response range of most stimuli continuously increases until it reaches its maximum. The minimum range is at a resolution of 1PPD. A closer investigation of the stimuli (Fig. 2.2 and the range (Fig. 3.3) suggests a relationship between the frequency, i.e. how often regions of an image alternate different intensities, and the range. Stimuli with higher frequencies, such as WE-circular1 and Checkerboard-0.938 (Fig. 2.2b and Fig. 2.2f), seem to display a higher range (Fig. 3.3). In conclusion, the ranges of model responses show similar growth and shrink patterns. Yet, the absolute range as well as the degree of change varies to a large extent and is dependent on the underlying stimulus.

The center of each distribution is close to zero. Furthermore, the distribution appears very symmetric. This is true throughout model responses of all resolutions and stimuli (Fig. 3.1 and Fig. 3.2). Consequently, the average value must be close to zero as well, because for each positive intensity value, a negative intensity value of nearly the same magnitude is supposed to exist. These values balance each other out, leaving the value of the center, which is zero. If that was not the case, the distribution would not be symmetric. It is common to express the average as the mean or the median. As expounded, both are expected to be close to zero throughout all underlying stimuli and resolutions. This finding is especially interesting because it shows that a model responses' property remains relatively persistent even if the model response resolution changes. To understand how model response properties change with altering stimulus resolution, it is also crucial to assess how they do not change. A more detailed analysis demonstrates that the model responses' mean and the median do not change significantly, despite their difference in resolution and underlying stimuli. The mean varies between -0.2 and 0.3; much as the median, which fluctuates between -0.4 and 0.3 (Fig.3.4a and Fig.??). Those values are considered to be close to zero; and their variations are rather small given an intensity value range, which is in some cases larger than 100 (Fig.3.3). This finding confirms the visual assumption that the center of the distribution is close to zero. Regardless of the small scale on which the mean and median change, it is noteworthy how they change: Except for the model response of Checkerboard-0.938, the mean and median of all other model responses converge towards zero with an increase in resolution. The detailed change depends on the underlying stimulus. While the mean and median of some model responses start to asymptote to zero from a negative value, others asymptote to zero, starting from a positive value. If the mean value of a particular model response is negative, the median value is as well; although in a different magnitude.

In addition to the previous findings, the distribution smooths out when the resolution decreases (Fig. 3.1 and Fig. 3.2). The directional changes of the distribution's graph decrease with decreasing size; the local minima and maxima of the graph converge on the y-axis until they finally merge into a monotonically decreasing or increasing function. The pattern becomes particularly visible for the distributions of WE-circular1, Todorovic-in-large, Todorovic-equal and SBC-large (Fig. 3.2a, Fig. 3.2b, Fig. 3.2c and Fig. 3.2d). Model responses of these stimuli with resolutions between 64PPD and 8PPD show especially much variability that gets lost when the size decreases. In summary, the histogram revealed several patterns that indicate how the model response of a stimulus changes (or does not change) when the respective resolution decreases. However, it became apparent that even though these patterns apply to most model responses, the exact nature of change depends on the individual stimulus.

Analyzing how model responses change with changing resolution yielded key findings, each of which is dependent on the stimulus and comes along with some exceptions. First, the model response range shows similar grow and shrink patterns throughout all stimuli. Low resolutions tend to contain low ranges, however that relation is non-linear. However, SBC-large differs from that general norm. Even though the model response reaches its maximum range at 8PPD like the other model responses, its minimum is reached at 64PPD. Second, the model response distributions converge to a median and mean of zero with an increasing resolution. However, the average is very close to zero for all stimuli, irrespective of their resolution. The model response of Checkerboard-0.938 is an outlier, because neither the mean nor the median of its model response converge to zero. Instead, they show a relatively stable value around 0.1 for all resolutions. Finally, all distributions are symmetric. The only exception to that are the distributions of Checkerboard-0.938s model responses (compare Fig.3.2e), which display slight asymmetries for resolutions of 16PPD, 32PPD and 64PPD. In contrast to the other distributions, it shows an asymmetric, sharp-edge peak and a larger area on the right than to the left of the center.

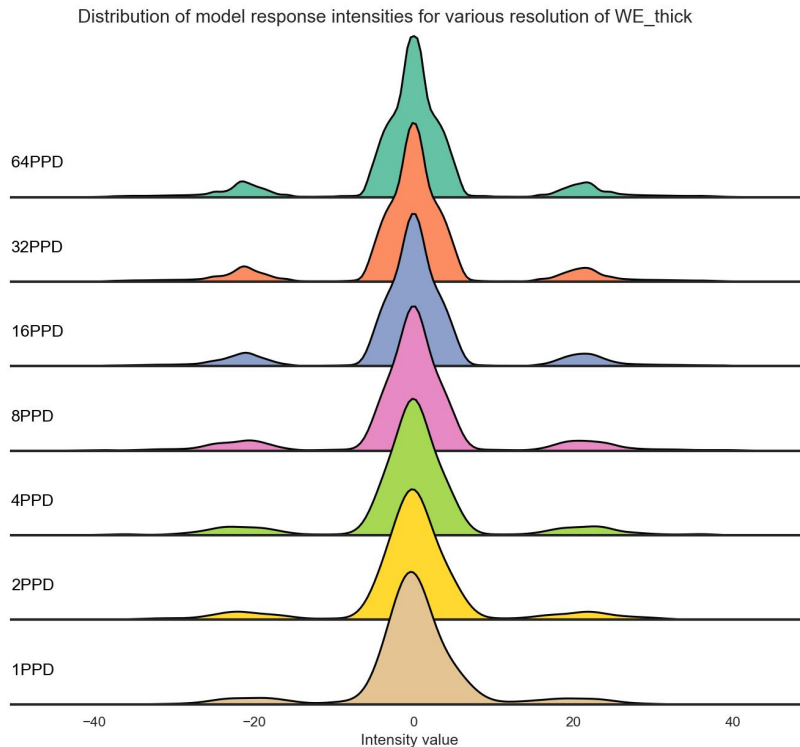


Figure 3.1: Each distribution represents model response intensity values for different resolutions of WE-thick. The resolution is shown on the left side of the distribution. The x-axis represents intensity values. The y-axis shows the relative occurrence of these intensity values. Each distribution is normalized such that their total area is equal. To display the distribution, the kernel density estimation (KDE) is used. The KDE is a widely used method to approximate the probability density function of a given dataset and plot its distribution[28].

### 3.2 Power spectrum analysis

The power spectrum analysis has produced two main findings: First, the total power of the model response increases with increasing resolution, reaches its maximum quickly and remains almost unal-

tered when the resolution is increased even more. Second, the total power of only the low frequencies stays very stable throughout various resolutions and stimuli.

The total power of an image is calculated as the sum of the normalized power spectrum of that image. For the assessed stimuli and resolutions, the total power of each model response remains very stable between 64PPD and 8PPD for most stimuli and decreases with decreasing resolution between 8PPD and 1 PPD (Fig. 3.5a). A more detailed analysis shows that the maximum total power of a model response is reached at 16PPD or 8PPD for each stimulus. However, the difference between the total power maximum at 16PPD and the total power at 32PPD or 16PPD is very small and invisible to the human eye at the line graph (Fig. 3.5a). The magnitude of the model response total power varies significantly and depends on the underlying stimuli: WE-circular1 has its largest power spectrum at 64PPD (234). Checkerboard-0.938 displays its largest total power at 8PPD with a magnitude of 70. The model responses of all stimuli have their minimum total power at a resolution of 1PPD. Yet, the magnitude of the minimum varies between a magnitude of 134 (SBC-large) and 14 (Checkerboard-0.938). The comparison of total power is helpful, but it comes with limitations. The highest image frequency that one can recover with the Discrete Fourier Transform depends on the sampling rate, i.e., the resolution of the image. This phenomenon is described by the Nyquist limit. The maximum frequency is dependent on the sampling rate of an image, i.e. the pixels per degree. Hence, the power spectrum of model responses with lower resolutions cannot recover the high frequency components of an image. Hence, if one compares the power spectra of model responses with different resolutions, one compares different frequency components of an image[29].

To compare the same frequency spectrum among different resolutions, one can use the frequency components that are present in all resolutions. Hence, the frequency components of the lowest resolution. The maximum frequency of the lowest resolution amounts to 0.5 which is half the sampling rate of 1PPD[29]. The comparison of the total power of the low model response frequency components up to 0.5 (TPLMRFC) shows that the model responses of almost all stimuli remain stable between 64PPD and 8PPD. The deviation between 64PPD, 32PPD and 16PPD is very small for all model responses. Even the deviation between 64PPD and 2PPD is tiny for most of them. The magnitude of the TPLMRFC is very dependent on the underlying stimulus: Todorovic-in-large has the highest TPLMRFC (134) and Checkerboard-0.938 the lowest (14).

In summary, the total power of a model response and the TPLMRFC remain very stable between 64PPD and 8PPD. However, the steadiness of the TPLMRFC appears to be little altered between 8PPD and 2PPD as well (compare Todorovic-in-large, WE-thick and SBC-large in Fig 3.5b). In, addition to these general patterns, a closer assessment of the total power suggests a link to the range (Fig. 3.3). Both, the range and the total power peak at 16PPD or 8PPD. From that peak, both decrease with decreasing resolution; and remain relatively steady with increasing resolution. Even though, the pattern of TPLMRFC appears similar to the range as well, it is different from range and total power in that it remains on a similar magnitude between 64PPD and 4PPD for many stimuli before it notably starts decreasing.

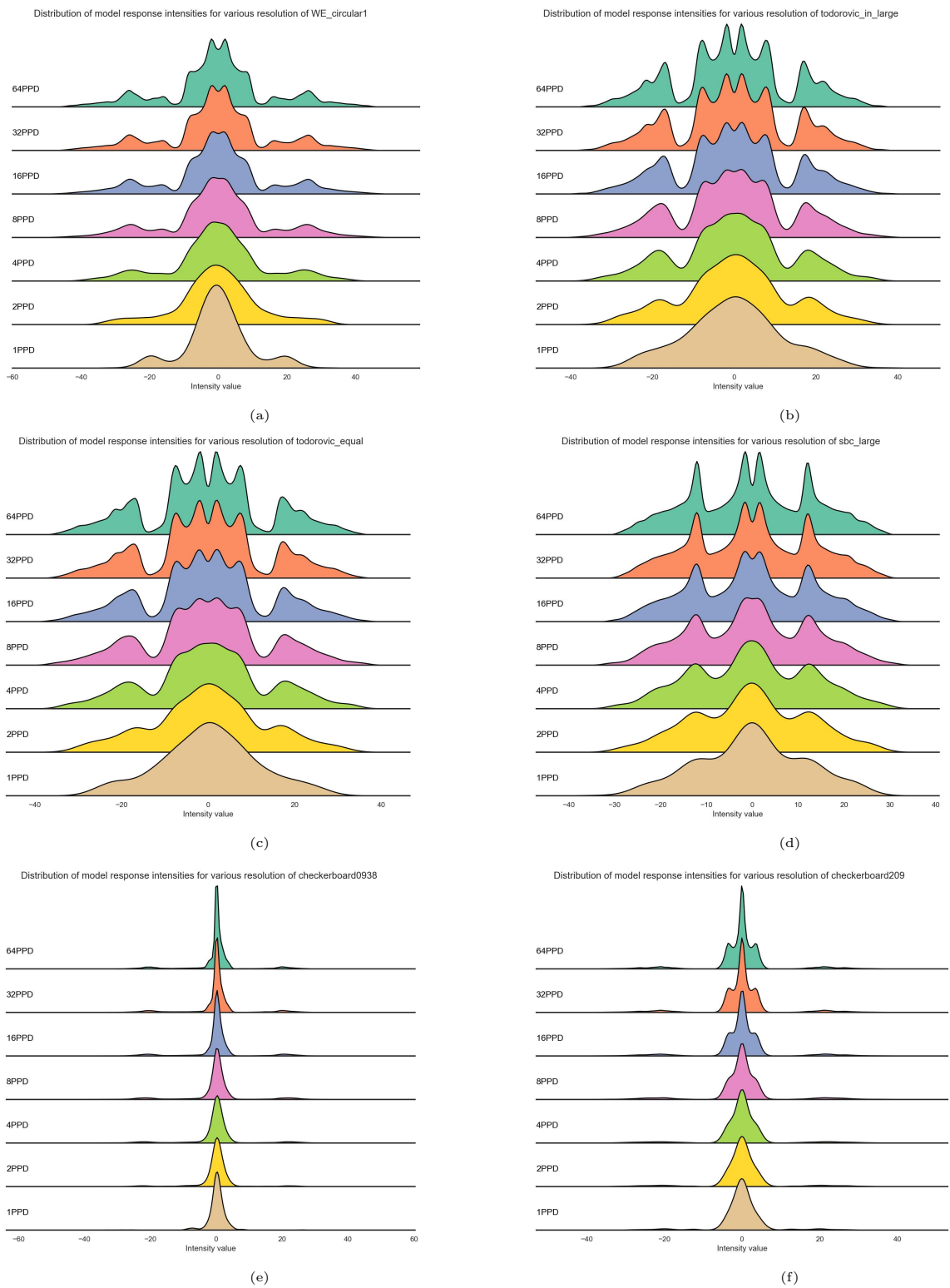


Figure 3.2

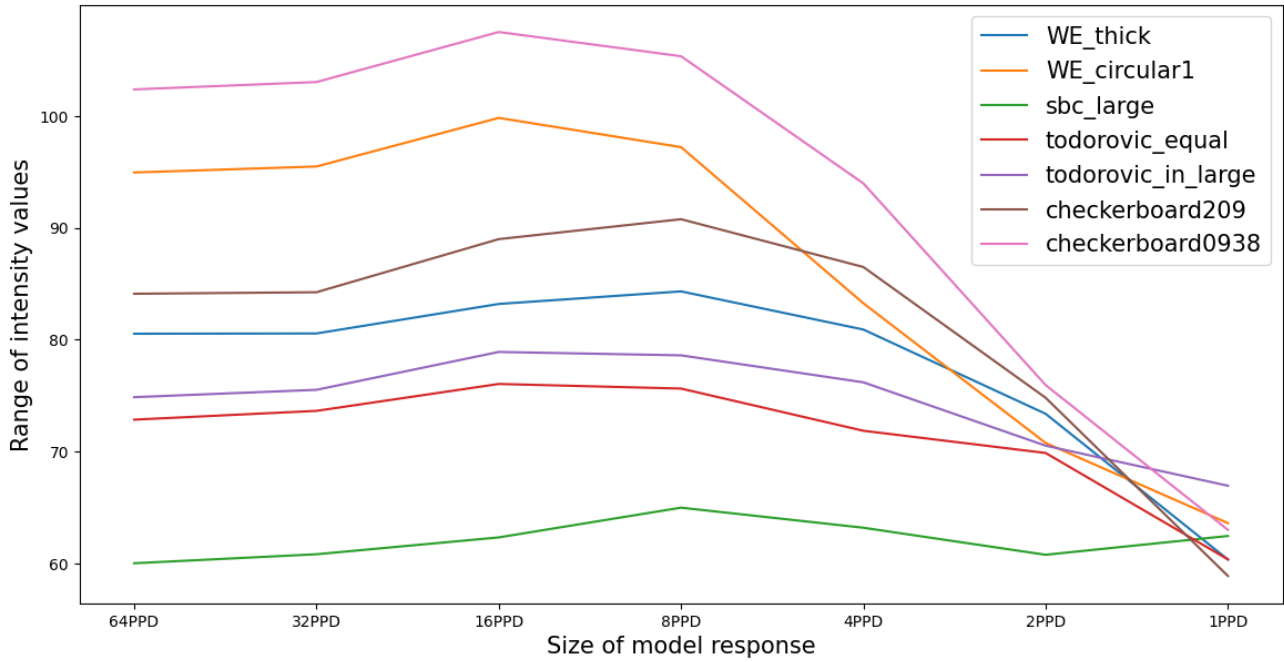
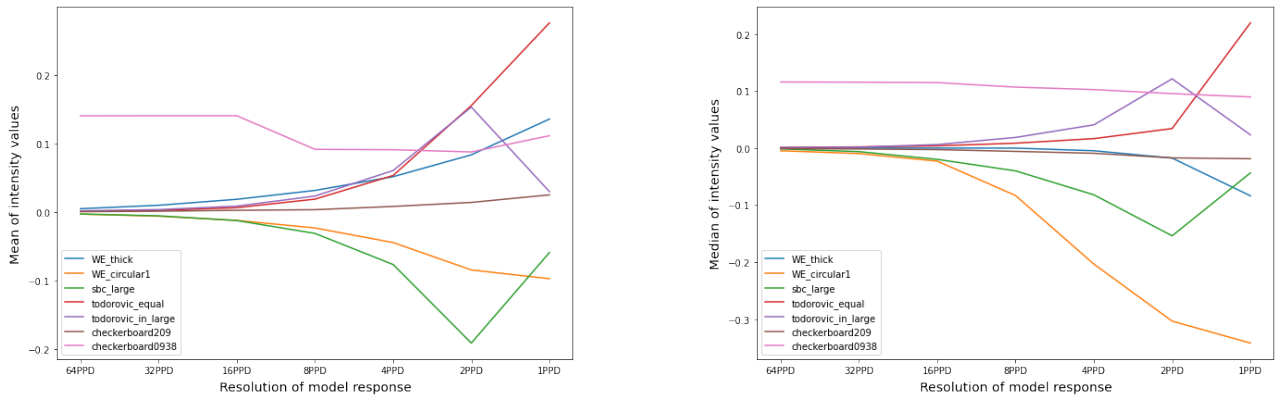


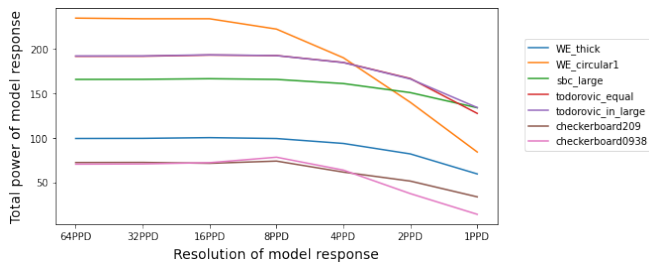
Figure 3.3: The figure shows the range of intensity values for model responses of different stimuli and resolutions. Given the minimum intensity value ( $min$ ) and maximum intensity value ( $max$ ) of a model response, the range was computed as  $max - min$



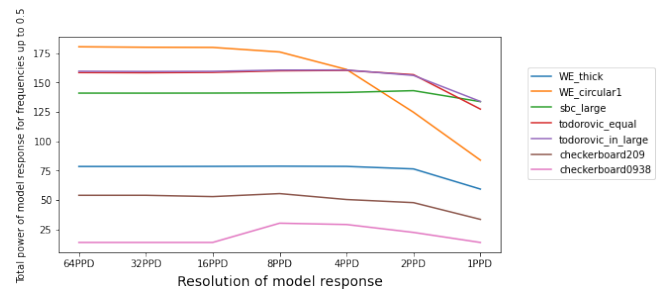
(a) The figure shows the mean intensity values for model responses of different stimuli and resolutions. The mean intensity value is calculated as the sum of all intensity values divided by the number of intensity values.

(b) The figure shows the median intensity values for model responses of different stimuli and resolutions.

Figure 3.4



(a) The model response total power of various resolutions and stimuli. The total power is calculated as the sum of power spectrum. The higher the resolution, the more frequency components can be considered.



(b) The model response power of frequencies up to 0.5 for multiple resolutions and stimuli. The total power of these frequencies is calculated as the sum of the power spectrum associated to frequencies up to 0.5.

Figure 3.5



## 4

# Discussion

The execution of multiscale spatial filtering models, that respond with a prediction of the perceived brightness of a stimulus (model response), is a time-consuming process (Chapter 1). The computations and, hence, the execution time are depended on the resolution of the stimulus (Chapter 2). The higher the resolution, the more computations, the slower the model execution process and vice versa. Consequently, one can reduce the model execution time by decreasing the resolution of the stimulus (Fig. 1.3). The reduction entails changes of the associated model response. Analyzing these changes is crucial for anyone that aims to speed up the model execution by reducing the stimulus resolution. To assess the model response changes, seven stimuli are used (Fig. 2.2). Each stimulus is created in various resolutions. The FLODOG model is applied to all variations of the stimulus. The yielded model responses are of different resolutions as well. To assess their (dis)similarity and understand how they change, the intensity distributions and power spectra of the model responses are compared (Fig 2.1). Indeed, the model responses of a stimulus vary significantly, depending on their resolution (Chapter 3). Even though detailed change depends on the stimulus, some common patterns are observed throughout almost all stimuli:

1. The range of intensity values change depending on the underlying stimulus.
2. The average intensity value converges to 0 with increasing model response resolution.
3. The distribution of intensity values smooths out when the resolution decreases.
4. Finding of power spectrum analysis.

The intensity range of model responses changes when the resolution of the underlying stimulus changes (Section 3.1). In particular, the range increases with increasing resolution until 8PPD or 16PPD, depending on the stimulus. For resolutions larger than that, the range starts to decrease again mildly. However, the detailed magnitude of change depends on the stimulus. One hypothesis would be that larger model responses contain a higher range of intensity values than smaller model responses. The reason for that is that larger model responses contain more pixels. Hence, they have a higher resolution and, in theory, the ability to display more details. This assumption was false for the model responses that were evaluated: The range of model responses increases between 1PPD and 8PPD and decreases again at either at 8PPD or 16PPD (Fig.3.3). Resolution and range do not correspond linearly. A closer investigation of the stimuli and the range of model responses suggests a link between the stimulus frequency and the model response range (Fig.2.2 and Fig 3.3): Checkerboard.0938 and

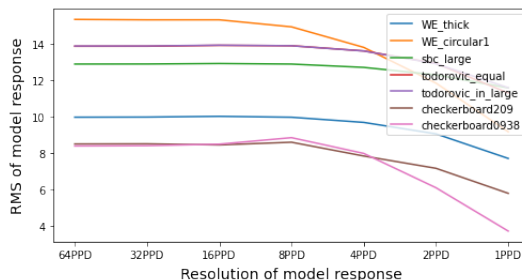


Figure 4.1: The figure shows the root-mean-square contrast (RMS) of model responses for different underlying stimuli and resolutions. The RMS is calculated according to Peli [31]

WE-circular1 alternate more often between black and white than for example SBC-Large or Todorovic-equal. Simultaneously, the stimuli with higher frequency of alternations, show higher ranges. This seems to be a systematic pattern. The pattern illustrates very well that the detailed change of the model response is dependent on the details of this stimulus itself; in this case, the different stimulus frequencies. In search for a potential explanation for that relationship, I did not find a direct link between the range of model responses and the frequency, but I found one between contrast and range. In fact, contrast, the difference of luminance, can be quantified by the range of brightness values[30]. A quick analysis of the 32PPD model responses has shown that the root-mean-square contrast(RMS) and the range show similar line graph patterns(Fig. 4.1 and Fig. 3.3). That seems reasonable: Contrast is defined as the difference in luminance. A higher range allows for a higher difference in luminance. Nonetheless, one can clearly see that the ranking of stimuli magnitude differs between the measure, i.e., the underlying stimulus of the model response with the highest range does not have the highest RMS.

The median and the mean of model responses converge to zero when the model response resolution increases (Fig. 3.4a and Fig. ??). This is true for the model responses of most examined stimuli. The center of the model response intensity distribution is close to zero and appears to be symmetric. The mean and median deviate no further than 0.4 from the expected value of zero. A general pattern is that smaller model responses deviate further from 0 than larger model responses. An assumption, why that deviation might occur, could be that the mean of smaller model responses is more susceptible to be influenced by outliers, because it contains fewer pixels. The influence of such outliers on the mean could have been reduced with increasing resolution, i.e. more pixels, which why the mean converges closer to zero accordingly. This assumption is false, because the median, which is not susceptible to outliers, shows a similar pattern of convergence with increasing resolution.

Moreover, the distribution smooths out when the resolution of the model response changes (Fig. 3.1 and Fig. 3.2). The smaller the resolution of a model response, the less variability in the distribution graph. One can interpret that as losing detail in the intensities, i.e., intensity values that are close to each other occur with approximately the same magnitude. Hence, all model response images are becoming more and more similar in intensity values.

With decreasing model response resolution, high-frequency information is lost. This loss is inherent by design of frequency-based image representation (section 3.2). However, the total power of the of model responses remain very stable between 64PPD and 8PPD (Fig. 3.5a). The total power decreases notably between 8PPD and 1PPD. This change cannot be surprising: The total would not differ if it was the same image. Furthermore, one could reason that with the loss of high-frequency information, one loses the amplitudes of those frequencies. Consequently, the total power of high resolution model

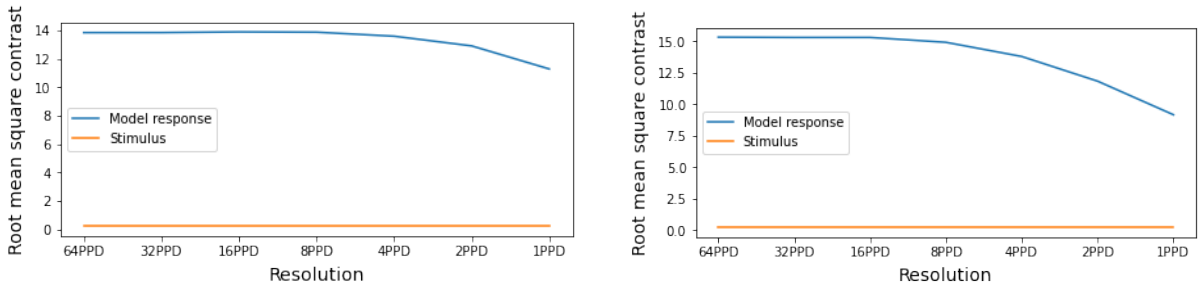
responses should be higher than the total power of low resolution model responses. That reasoning is false. In fact, the model response power for all model responses peaks at either 8PPD or 16PPD, even though high-spatial frequency information is lost. Moreover, one could make the assumption that despite the loss of high-spatial frequency information, low spatial frequency information of all model responses are the same, irrespective of their resolution. This is not true. Comparing the total power of the low model response frequency components (TPLMRFC)(Fig. 3.5b) shows that the magnitude of TPMRFC differs between resolutions. That show that the model responses differ in low frequencies as well. However, the TPMRFC does not change as notably as the total power, which suggests that with decreasing resolution, more high than low-frequency information is lost. In addition, the total power and the range of model responses, show a similar line graph pattern. However, even more obvious is the link between RMS and the total power of a model response. Their graphs seem to be identical (Fig. 3.5a) and Fig. 3.3). Indeed, the literature suggests that the squared RMS and the total power (energy) are equal[32].

In some cases, scientist rely on the range[3] and the average[1] to be consistent when comparing the brightness of two patches<sup>1</sup>, e.g., comparing a WE-thick model response patch with an SBC-large model response patch. Scientists use the range and the average to normalize the intensity values of the patches. This normalization step is indispensable to compare the intensity of the patches. It is easy to illustrate that: A patch with an intensity of -1 has a different significance for a stimulus with a range of intensity values between -1 and 1 than for a stimulus with a scale between -100 and 100. The range and the average (mean median) of a model response change notably, when the size of the stimulus changes. Anyone who decreases any stimulus, to improve the model execution speed, should be aware that they cannot neither rely on a steady range nor a steady average of intensity values for the associated model response. Both, the range and the average of the model response change with decreasing resolution. However, the absolute magnitude of change varies clearly noticeable. Suppose, the resolution is decreased from 32PPD to 16PPD. The difference in mean and median would become noticeable in the second or third decimal place. In contrast, the difference in range varies between 2 and 5. If one uses the model response range or the average for normalization purposes, one has to be aware of the change caused by decreasing the resolution of the stimulus. Instead of using the range to normalize, one could consider using the total power. Contrast and range are related. Range and total power are related. The absolute magnitude of total power change between 32PPD and 16PPD is less than 1.5. Hence, if one wants to decrease the stimulus and normalize the brightness of the patches, a method that uses the contrast for normalization, could be more reliable than normalizing with the range. Moreover, the power of lower frequencies changes less with decreasing resolution than the total power of low model response frequencies, which suggests that one can consider the use of the lower frequencies even if the resolution of the model response decreases from 32PPD to 16PPD. Despite all these considerations, one has to be aware that the exact changes of the model response are based on the underlying stimulus. Consequently, if one wants to decrease the stimulus "only" from 32PPD to 16PPD, it is crucial to keep any model response property in check that one wants to utilize later on.

The total power of the model response changes with decreasing resolution. In addition, there is an equality of power and contrast. Hence, the model response contrast changes, with changing resolution as well (Fig. 4.1). That is especially interesting, given the fact that the contrast of the input stimulus does not change with changing resolution (Fig. 4.2a and Fig. 4.2b). An obvious assumption would be that the contrast of the model response changes because the perceived contrast changes. That

---

<sup>1</sup>Patches are the areas of a stimulus that are of identical luminance, but that humans perceive as differently bright.



(a) The root-mean-square contrast of the Todorovic-equal stimulus; and the FLODOG model response of Todorovic-equal.

(b) The root-mean-square contrast of the WE-circular1 stimulus; and the FLODOG model response of WE-circular1.

Figure 4.2

would suggest that the FLODOG model does not only model perceived brightness, but also perceived contrast. Indeed, [p.1641]Robinson, Hammon, and Sa [1] mentions in his results that the model predicts contrasts according to some peoples' perception. The FLODOG model aims to model early human vision. Contrast and brightness perception are both expected to occur in the early stages of human vision[1]. Hence, it is not surprising if the FLODOG model were to predict contrast as well. Another finding is that all stimuli have the same contrast. However, the contrast of their model response changes.

The aforementioned patterns come along with some exceptions: The SBC-large model response shows model response ranges that deviate from model responses of other stimuli. It shows a higher range of intensity values at 1PPD than at 64PPD, which is contradictory to the patterns of other model response (Fig. 3.3). The average intensity for the model responses of Checkerboard-0.938 does not converge to zero like the other model responses (Fig. ?? and Fig. 3.4a). The average remains at an almost constant value close to zero, irrespective of the resolution. Moreover, the model response intensity distribution of Checkerboard-0.938s deviates from other model responses due to its slight asymmetry (Fig. 3.2e). The distribution tends to occupy more space on the right than to the left of the center. The center of the distribution is close to zero. That means, the area on the right of the center represents the positive intensity values. This circumstance allows the assumption that the model response of Checkerboard-0.938 contains more pixels with high intensity values than pixels with low intensity values. However, the visual impression of Fig.4.3 does not seem to support this assumption at first glance. It rather suggests the opposite, because the model response seems to contain a surplus of negative intensity values. This impression is caused by the corners of the 32PPD checkerboard as well as their immediate surrounding. They appear bluer and hence smaller than other parts of the model response. In addition to that, the 1PPD model response contains more rectangles with negative intensity values. However, this suggestion might be a misconception. In contrary, these corners could be the reason why the distribution of Checkerboard-0.938 shows an excess of positive values near the center. This might seem counterintuitive, but it becomes clearer, when one takes a closer look at Fig.3.4a and Fig.3.4a. The mean and median of most model responses have the tendency to converge towards zero. Hence, the model response contains intensity values that are balanced between negative and positive intensity values. For the model response of Checkerboard-0.938 to be balanced, it needs to contain high intensity values to counteract the imbalance caused by the corners. Indeed, a visual inspection of the 1PPD model response in Fig.4.3 leaves the spectator with the impression of a reddish background, which is an indication of a slight surplus of positive intensity values.

How do model responses change when the resolution of the underlying stimulus changes? Any user

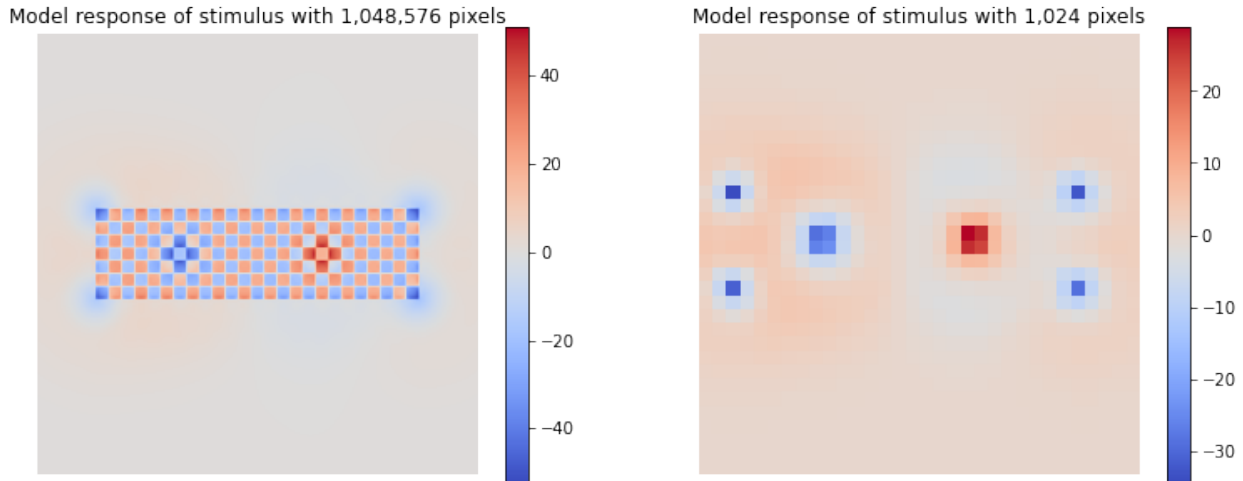


Figure 4.3: Left: Response of the FLODOG model applied to Checkerboard-0938 effect with 1,048,576 pixels. Right: The same model applied to the same stimulus with 1,024 pixels. The color bar shows the intensity value of the corresponding model responses. Lower intensity values, represent less brightness and vice versa.

of the model, who wants to decrease the stimulus resolution to increase the model execution time, has to be aware that properties of the model response will change and that those changes are non-trivial to predict because the degree and magnitude of change are very dependent on the underlying stimulus. If those changes are acceptable for their use case, can only be answered by the user. Despite those changes, some properties are more robust than others. The FLODOG model seems to contain an inherent mechanism that causes the average to converge to zero and maintain an even range. Even the mean and median of low resolutions remain close to zero. The model response range behaves less predictable. Its total value is very dependent on the stimulus. It neither falls nor rises reliably with increasing or decreasing resolution. Some properties could have resisted change completely. The power of low frequencies could have stayed exactly the same, because with decreasing resolution, only high frequency information is lost. But in fact, also the low frequency information of the model response change with changing resolution. The fact that some properties change less than others, seems to reveal more information about the inherent behavior of the models. Maybe these findings are self-evident for the originators or frequent users. Nonetheless, to my knowledge, these properties have not been mentioned explicitly in any of the original papers. For someone who is starting to use these models, making these model behaviors explicit, can only ever be a contribution.

# List of Figures

1.1	White's Effect and the associated FLODOG model response . . . . .	7
1.2	Spatial filter with 3x3 neighborhood . . . . .	8
1.3	Execution time of the FLODOG model . . . . .	8
1.4	Model response of White's Effect: Comparison of different resolutions . . . . .	9
2.1	Methodological approach: A flowchart . . . . .	11
2.2	All stimuli and their corresponding model response . . . . .	13
2.3	Schematic of visual angle . . . . .	14
2.4	Dog and OdoG filter . . . . .	15
2.5	Intensity distribution of WE-thick . . . . .	16
3.1	Model response intensity distribution of various resolution of WE-thick . . . . .	20
3.2	Model response intensity distribution of various resolution of multiple stimuli . . . . .	22
3.3	Range of intensity values . . . . .	23
3.4	Average of intensity values . . . . .	23
3.5	Power spectrum of model responses . . . . .	24
4.1	Contrast of model responses . . . . .	26
4.2	Contrast of stimulus compared to contrast of model response . . . . .	28
4.3	Checkerboard-0.938: Model responses of different resolutions . . . . .	29

# Bibliography

- [1] A. E. Robinson, P. S. Hammon, and V. R. de Sa, “Explaining brightness illusions using spatial filtering and local response normalization,” *Vision research*, vol. 47, no. 12, pp. 1631–1644, 2007.
- [2] F. Kingdom, *Simultaneous contrast: The legacies of hering and helmholtz*, 1997.
- [3] B. Blakeslee and M. E. McCourt, “A multiscale spatial filtering account of the white effect, simultaneous brightness contrast and grating induction,” *Vision research*, vol. 39, no. 26, pp. 4361–4377, 1999.
- [4] M. White, “A new effect of pattern on perceived lightness,” *Perception*, vol. 8, no. 4, pp. 413–416, 1979.
- [5] R. F. Murray, “Lightness perception in complex scenes,” *Annual Review of Vision Science*, vol. 7, pp. 417–436, 2021.
- [6] B. Blakeslee and M. E. McCourt, “Similar mechanisms underlie simultaneous brightness contrast and grating induction,” *Vision research*, vol. 37, no. 20, pp. 2849–2869, 1997.
- [7] M. J. McMahon, O. S. Packer, and D. M. Dacey, “The classical receptive field surround of primate parasol ganglion cells is mediated primarily by a non-gabaergic pathway,” *Journal of Neuroscience*, vol. 24, no. 15, pp. 3736–3745, 2004.
- [8] R. C. Gonzalez and R. E. Woods, *Digital image processing*, 4th ed. Pearson Deutschland, 2017.
- [9] E. Demidenko, *Mixed models: theory and applications with R*. John Wiley & Sons, 2013, p. 610.
- [10] Q. Bao and P. Guo, “Comparative studies on similarity measures for remote sensing image retrieval,” in *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, IEEE, vol. 1, 2004, pp. 1112–1116.
- [11] T. M. Lehmann, M. O. Güld, C. Thies, *et al.*, “Content-based image retrieval in medical applications,” *Methods of information in medicine*, vol. 43, no. 04, pp. 354–361, 2004.
- [12] F. Rajam and S. Valli, “A survey on content based image retrieval,” *Life science journal*, vol. 10, no. 2, pp. 2475–2487, 2013.
- [13] K. Arun and H. P. Menon, “Content based medical image retrieval by combining rotation invariant contourlet features and fourier descriptors,” *International Journal of Recent Trends in Engineering*, vol. 2, no. 2, pp. 35–36, 2009.
- [14] Y. Wang, Q. Chen, and B. Zhang, “Image enhancement based on equal area dualistic sub-image histogram equalization method,” *IEEE transactions on Consumer Electronics*, vol. 45, no. 1, pp. 68–75, 1999.
- [15] A. Hyvärinen, J. Hurri, and P. O. Hoyer, *Natural image statistics: A probabilistic approach to early computational vision*. Springer Science & Business Media, 2009, vol. 39, pp. 33–34.

- [16] P. D. Howe, “White’s effect: Removing the junctions but preserving the strength of the illusion,” *Perception*, vol. 34, no. 5, pp. 557–564, 2005.
- [17] D. Todorović, “Lightness and junctions,” *Perception*, vol. 26, no. 4, pp. 379–394, 1997.
- [18] L. Pessoa, G. Baratoff, H. Neumann, and D. Todorovic, “Lightness and junctions: Variations on white’s display,” *Investig. Ophthalmol. Vis. Sci.(Suppl.)*, vol. 39, S159, 1998.
- [19] B. Blakeslee and M. E. McCourt, “A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization,” *Vision research*, vol. 44, no. 21, pp. 2483–2503, 2004.
- [20] S. Mathôt. “Degrees of visual angle.” (), [Online]. Available: <https://osdoc.cogsci.nl/3.3/visualangle/>.
- [21] R. W. Rodieck, “Quantitative analysis of cat retinal ganglion cell response to visual stimuli,” *Vision research*, vol. 5, no. 12, pp. 583–601, 1965.
- [22] J. Vincent. “Creating and applying image filters.” (), [Online]. Available: [https://multyscale.readthedocs.io/en/latest/tutorials/demo\\_filters.html#Filter-types](https://multyscale.readthedocs.io/en/latest/tutorials/demo_filters.html#Filter-types).
- [23] C. Wang, J. P. Reese, H. Zhang, *et al.*, “Similarity-based visualization of large image collections,” *Information Visualization*, vol. 14, no. 3, pp. 183–203, 2015.
- [24] K. R. Castleman, *Digital Image Processing*. Prentice Hall Professional Technical Reference, 1979, pp. 71–73.
- [25] B. Jaehne, “Image representation,” *Digital Image Processing*, 2005.
- [26] L. Schmittwilken, M. Matko, G. Aguilar, and M. Maertens, *Stimuli(v.1.1)[python package]*, 2022. [Online]. Available: <https://github.com/computational-psychology/stimuli>.
- [27] J. Vincent, *Multyscalescale(v.1.0)[python package]*, 2022. [Online]. Available: <https://github.com/computational-psychology/multyscale>.
- [28] Y.-C. Chen, “A tutorial on kernel density estimation and recent advances,” *Biostatistics & Epidemiology*, vol. 1, no. 1, pp. 161–187, 2017.
- [29] E. W. Weisstein, *Nyquist frequency*. [Online]. Available: <https://mathworld.wolfram.com/NyquistFrequency.html>.
- [30] M. Nixon and A. Aguado, *Feature extraction and image processing for computer vision*. Academic press, 2019, p. 84.
- [31] E. Peli, “Contrast in complex images,” *JOSA A*, vol. 7, no. 10, pp. 2032–2040, 1990.
- [32] D. G. Pelli and P. Bex, “Measuring contrast sensitivity,” *Vision research*, vol. 90, pp. 10–14, 2013.