The Influence of Noise on Human Edge Perception in Natural Images

Jarl-Sebastian Sørensen

402177

submitted 31st March 2023

Bachelor Thesis

Computer Science (B.Sc.)

Technische Universität Berlin

Faculty IV – Electrical Engineering and Computer Science

Institute of Computer Engineering and Microelectronics

Dept. Computational Psychology

First examiner:

Prof. Dr. Marianne Maertens

Second examiner:

Prof. Dr.-Ing. Olaf Hellwich

Supervisor:

Lynn Schmittwilken

Declaration of Academic Integrity

I hereby declare that this thesis has been written without unauthorized outside help and that no sources and aids other than those indicated have been used.

Berlin, 31. March 2023

Jarl-Sebastian Sørensen

Abstract

Edges are important visual features of our environment and it is widely believed that extracting edges from the visual input is an early step in human visual processing. Evidence has accumulated that edge detection may be mediated by a narrow spatial scale around 3 cpd. However, most evidence has been gathered with isolated edges and hence simplified stimuli. As it is an open question to which extent these simplified stimuli unveil the inner workings of the visual system, complementary research using natural stimuli is necessary. Since there is no standard approach for testing human edge perception in natural images, this thesis has two aims: (1) We have developed and evaluated a method for testing human edge perception in natural scenes, and (2) we have used a noise-masking paradigm to investigate whether human edge perception in natural scenes is similarly affected by noises that interfere with spatial frequency contents around 3 cpd as it is the case for isolated edges. The basis of our approach is a contour segmentation task, in which participants are instructed to segment the outlines of the image contents in a natural image using a self-created segmentation tool. To quantify human performance and to measure the quality of our approach, we compare the resulting segmentation maps with a ground truth using a similarity heuristic. Assessing our method, it seemed to be valid and effective to test edge perception in natural images. Consistent with prior literature, we found that also for more naturalistic stimuli, edge perception deteriorates most when interfering with image contents at a spatial scale of 3 cpd.

Zusammenfassung

Kanten sind wichtige visuelle Merkmale unserer Umwelt und es wird allgemein angenommen, dass die Extraktion von Kanten aus dem visuellen Input ein früher Schritt in der menschlichen visuellen Verarbeitung ist. Es hat sich gezeigt, dass die Erkennung von Kanten durch eine Raumfrequenz um 3 cpd vermittelt wird. Die wissenschaftliche Grundlage dafür wurde jedoch größtenteils mit isolierten Kanten und daher mit vereinfachten Stimuli geschaffen. Da es eine offene Frage ist, inwiefern diese vereinfachten Stimuli das Innenleben des visuellen Systems enthüllen, ist ergänzende Forschung mit natürlichen Stimuli notwendig. Bis heute gibt es jedoch keinen Standardansatz, um die menschliche Kantenwahrnehmung in natürlichen Bildern zu testen. Daher verfolgt diese Arbeit zwei Ziele: (1) Wir haben eine Methode zum Testen der menschlichen Kantenwahrnehmung in natürlichen Szenen entwickelt und evaluiert und (2) wir haben ein "noise-masking" Paradigma verwendet, um zu untersuchen, ob die menschliche Kantenwahrnehmung in natürlichen Szenen in ähnlicher Weise durch Bildrauschen beeinträchtigt wird, das einem räumlichen Frequenzgehalt von 3 cpd entspricht, wie es bei isolierten Kanten der Fall ist. Die Grundlage unseres Ansatzes ist eine Konturensegmentierungsaufgabe, bei der die Teilnehmer angewiesen werden, die Umrisse der Bildinhalte in einem natürlichen Bild mit Hilfe eines selbst erstellten Segmentierungswerkzeugs zu segmentieren. Um die Kontursegmentierunsgsleistung zu quantifizieren, aber auch um die Qualität unseres Ansatzes zu messen, vergleichen wir die resultierenden Segmentierungen mit einer Grundwahrheit unter Verwendung einer Ähnlichkeitsheuristik. Beim Testen unserer Methode haben wir festgestellt, dass unser Ansatz valide und effektiv scheint, um die Kantenwahrnehmung in natürlichen Bildern zu testen. In Übereinstimmung mit früherer Literatur fanden wir heraus, dass die Kantenwahrnehmung auch bei natürlichen Stimuli am stärksten verschlechtert wird, wenn der Bildinhalt in einer Raumfrequenz von 3 cpd beeinträchtigt wird.

Table of Contents

1 Introduction

1.1 Edges

Edges carry information about object and surface boundaries in images. An edge is a luminance step in space that is elongated in one direction. Edges are detected by neurons in the human visual system and edge detection is widely believed to be a first step in many visual processing tasks, such as surface representation (Salmela & Laurinen, 2005) or object recognition (Biederman, 1987). Furthermore, a primal sketch as introduced by Marr (1976) is often theorized to be constructed at the start of human visual processing. The primal sketch is a description of the basic elements of an image, usually containing edges. For these reasons, a better understanding of edge detection in the human visual system could be key to gaining insight into higher-level functionalities in human vision. Therefore, we want to further study human edge perception in this thesis. Before we explain our approach, we will introduce some general concepts on human edge perception and the underlying physiology that are needed to further introduce and motivate our research.

The signals coming from the retina of the eye reach an early stage of visual processing in the primary visual cortex (V1). There, neurons labeled simple cells process the retinal information by responding selectively to stimuli of a certain spatial frequency (also called spatial scale) and orientation. Spatial frequency refers to the number of pairs of bars in an image within a given distance on the retina and is usually expressed in cycles per degree (cpd) (Figure 1)[1].

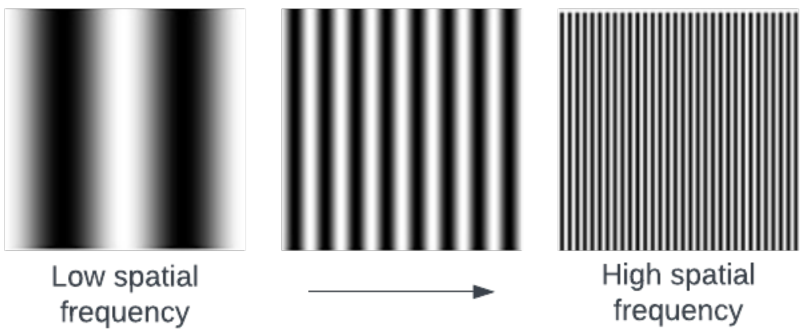

Figure 1: Visual representation of spatial frequency. From left to right the spatial frequency increases. https://www.rochester.edu/newscenter/microscopic-eye-movements-affect-how-we-see-contrast-358802

[1] http://www.psy.vanderbilt.edu/courses/hon185/SpatialFrequency/SpatialFrequency.html

Hubel and Wiesel (1962) first discovered these simple cells in the primary visual cortex of a cat and mapped their receptive fields. The receptive field of a cell in the visual system is the area of the retina that influences the response of the cell (Hubel & Wiesel, 1962). The receptive field of simple cells was shown to consist of parallel inhibitory and excitatory regions of approximately the same size. These regions are separated by parallel lines. If light hits the receptive field of a simple cell equally in both areas, the responses will cancel each other out, and hence the simple cell will not respond to light impulses larger than its receptive field. But, if an unregular pattern of light falls across the receptive field, the responses from the inhibitory and the excitatory areas will not be the same and will instead be integrated to form a response. Therefore, the response is strongest if light just hits the excitatory area of the receptive field. That light impulse would be of the same orientation as the simple cell's receptive field and its spatial scale would relate to the size of the receptive field (Elder & Sachs, 2004).

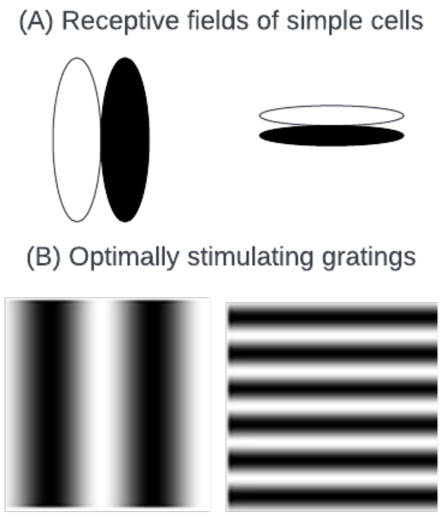

Figure 2: Visual representation of receptive fields (A) and respective optimal stimuli (B). In (A) the bright regions indicate the excitatory and the dark regions the inhibitory part. The scale and orientation of the stimuli match the scale and orientation of the receptive fields.

Hence, a simple cell responds strongest to stimuli with a certain orientation and spatial frequency, depending on the composition of its receptive field. For a visual representation see Figure 2. Thereby, simple cells could function as filters decomposing the visual input into features with different spatial properties (De Valois et al., 1982).

Some specific simple cells were additionally shown to respond strongest to stimuli representing edges (Hubel & Wiesel, 1962). Therefore, they were later speculated to account for edge detection and were hence called edge detectors (Tolhurst, 1972). Shapley and Tolhurst (1973) measured the spatial frequency response of presumed edge detectors in a psychophysical experiment. Their results showed that edge detectors seem to be most sensitive to stimuli with a spatial frequency of 3 cpd. The results of Schmittwilken and Maertens (2022b) from a noise-masking experiment support the results from Shapley and Tolhurst (1973). They also concluded that a relatively narrow spatial scale around 3 cpd plays a special role in extracting edge information from the visual input. Given the assumption that simple cells respond to oriented information with a specific spatial frequency, this implies that simple cells with a spatial frequency selectivity of 3 cpd account for edge detection.

They used noise to mask their stimuli because noise introduces random fluctuations to a signal that hinders an accurate perception of it. Different types of noise exist, often containing different frequency bands. Those different noises hinder only the perception of the frequencies in the signal that are also contained in the noise. Therefore, if noise is applied to an image the contents of the image in the frequency range of the noise are distorted and cannot be perceived anymore. That way, noise can be used to unveil the sensitivity of neurons to contents in certain spatial frequency ranges, e.g. the spatial frequency response of simple cells.

1.2 From Simple Stimuli to Natural Images

Many experiments use simple stimuli like sine wave gratings or bars and spots to study edge perception. While using stimuli like these has long been standard practice in vision research, they have also attracted criticism. The concern is that those stimuli might not be suitable to truly unveil the functioning of the human visual system (Olshausen & Field, 2005). This has given prominence to using natural images, meaning images of the real world. One argument for the use of natural images as stimuli is that the responses of neurons to these simple stimuli may not translate to natural images (Olshausen & Field, 2005; Touryan & Dan, 2001). As artificial stimuli are often designed around assumptions of how the responses from neurons should be, they may not contain features relevant to the response of a neuron (Touryan & Dan, 2001). Additionally, neuronal processing is largely nonlinear, and therefore responses are often poorly predicted by artificial stimuli alone (Felsen & Dan, 2005; Touryan & Dan, 2001). Examples of response properties uniquely revealed by natural images are the research by Felsen et al. (2005) and David et al. (2004). Olshausen and Field (2005) state that the only way to truly be able to predict the responses from a neuron is to measure the responses for every possible stimulus. As this is not possible, natural images could be used instead. The reason is that natural images match the real context the visual system processes, i.e. it is an ecologically valid test. This argument for the use of natural images is taken one step further by the efficient coding hypothesis, which holds that the purpose of early visual processing is to produce an efficient representation of the incoming visual signal (Simoncelli, 2003). This implies that because the sensory circuits evolved in a natural environment, they may be specifically tuned for the efficient coding of natural stimuli (Touryan & Dan, 2001). Natural stimuli have been shown to be far from random, instead, they possess distinct properties (Simoncelli & Olshausen, 2001). Therefore, if the efficient coding hypothesis holds true, natural stimuli could be essential to gain insight into the underlying mechanisms of the human visual system. Given the above reasoning, it is not clear whether insights on human edge perception that have been studied with simplified gratings might translate to natural stimuli. Hence, the finding of Shapley and Tolhurst (1973) and Schmittwilken and Maertens (2022b) that a narrow spatial scale of 3 cpd is most informative about edges does not necessarily translate to natural images. Therefore, complementary research with natural images is necessary.

1.3 Aim of Thesis

Edges are an important feature in our environment and edge detection is widely believed to be an early stage of human visual processing. Since edges are classically studied with simplified stimuli, it is not clear to which extent our knowledge about the underlying mechanisms of human edge perception translates to natural stimuli. Hence, we wanted to test edge perception in natural images to draw conclusions about which spatial frequency contents are most important for perceiving edges. However, while designing an experimental paradigm to investigate this question, we came to realize that testing human edge perception in natural images is a challenging task and that there is no standard approach for it yet. Hence, the aim of this thesis became two-fold. First, we develop an approach to investigate human edge perception in natural images. We test our approach using a similar noise-masking paradigm as Schmittwilken and Maertens (2022b), however, we exchange their simplified stimuli with a natural image. In a second step, we then use our approach to investigate whether spatial frequency contents around 3 cpd also play an important role for human edge perception in natural images.

1.4 Summary

In the experimental paradigm, the focus lies on contours (i.e. outlines of image contents) instead of edges. Observers are instructed to segment contours in a self-created segmentation tool. The resulting contour maps are then compared to a ground truth using a heuristic specified by Grigorescu et al. (2003), resulting in a score indicating the quality of the segmentation. We tested this paradigm in a laboratory environment. The basis for our stimuli are the images from the natural image dataset of Grigorescu et al. (2003). One natural image masked with different noises was presented over a range of contrasts. Through this, we hoped to find answers on the efficacy of the experimental design. We found that the use of ground truths for comparison is valid and that the segmentation quality measure seems to match our perception of segmentation quality. Image contrasts below 0.04 should not be used in our case and an error margin in the segmentation quality measure changes results only quantitatively. Further, preliminary results indicated that narrowband noise of 3 cpd and pink noise are most obstructing to contour detection in natural images. Finally, we discussed the implications of these results.

2 Methods

In this section, we outline our experimental design. We begin by proposing and discussing the experimental paradigm on edge perception in natural images in section 2.1. Then, we provide details on how we apply this experimental paradigm specifically to our test case in section 2.2.

2.1 General Considerations

Since it is not clear how to design an experiment investigating edge perception using natural images, we propose the following.

Instead of general edges, we focus on contours, i.e. the outlines of the image contents. The observers are instructed to segment the contours in a stimulus using a self-created segmentation tool. The resulting contour images are then compared to a ground truth with a similarity measure. We discuss the different parts of the experimental paradigm in detail in the following sections. We start by discussing a selection of natural images in section 2.1.1. We then explain the participant's task in detail in section 2.1.2. Finally, we delve into how this paradigm can be implemented in sections 2.1.3 and 2.1.4.

2.1.1 In Search of a Natural Image Database

Our goal is to compare contour segmentations of a stimulus to a ground truth, i.e. a human-labeled contour segmentation on the undistorted image. For this, it is necessary to have a dataset that contains natural images and their corresponding contour maps. We examined several datasets and ultimately selected one to utilize in the experiment. Natural image datasets with human-drawn contour maps are Grigorescu et al. (2003), Li et al. (2019), and Martin et al. (2001).

The Martin et al. (2001) dataset contains 300 grayscale and 300 color images of various sizes, with multiple contour maps for each image. Additionally, a segmentation tool is provided with it. However, a limitation of this dataset for our case is that the supplied contour maps solely consist of closed contours, i.e. contours that share the same start and end point. This is an issue for our purpose as it makes the segmentation very unintuitive, thereby increasing the time required for the task. An example from this dataset displaying the problem is shown in Figure 3 (A).

The dataset by Li et al. (2019) contains 1000 color images of different dimensions, each image has 5 corresponding contour maps. What makes this dataset unfit in our case is that the contour maps are just roughly edge aligned. This means that an edge present in the ground truth may not necessarily correspond to an edge present in the image, which is undesirable in our case. Figure 3 (B) shows an example from the dataset displaying the issue.

Figure 3: Example natural images from the different datasets with respective contour maps. In (A) it is visible that the nose and the fingers of the woman are not traced, because only closed contours could be segmented. In (B) the tracing of the dog is very rough and does not fit the actual contours in the image. In (C) one example of an animal and a human made object is listed.

We use the dataset of Grigorescu et al. (2003). It consists of 40 images with respective contour maps. Mostly, different animals in their natural environments are depicted with the addition of some human-made objects. The images are all grayscale and 512×512 px in size. Examples from the image dataset of Grigorescu et al. (2003) with their corresponding contour maps are listed in Figure 3 (C). Although we decided to utilize this specific dataset, it is worth noting that the other two datasets have the advantage of multiple contour maps per image, creating a more representative benchmark for human contour perception. Having selected a natural image dataset, we will discuss how to study edge detection in natural images in the next section.

2.1.2 In Search of a Task

The study of edge perception typically involves using simple stimuli designed to elicit a strong neural response from the visual system. These stimuli often contain a fixed number of edges in a specific direction, e.g. gratings or bars (See Figure 2). A common experimental paradigm using this type of stimulus is to present multiple of these stimuli to observers, who are then asked to detect the presence of edges or their direction. An example of such an experiment is the work by Schmittwilken and Maertens (2022b).

The same approach cannot be easily applied to natural stimuli, though, as it is debatable what constitutes an edge in a natural image (Heath et al., 1998), especially to a human observer. Therefore, we first need a clear definition of the term edge. This is not straightforward, because a definition must not necessarily be functional for a human observer. Figure 4 (A) shows edges as they are defined by a Canny edge detector. However, such a segmentation could not be created by a human. To address this problem, we want to introduce the concept of contours as defined by Grigorescu et al. (2003). They defined contours as a subset of edges specifying the outlines of the image contents, not including edges originating from textured regions (Figure 4 (B)). As we have a strong intuition for what contours represent in an image (i.e. the outlines of image contents), they are easily detected by an observer.

Still, it might not work if observers were just asked to indicate whether contours were present in a stimulus, as the contours may be visible under different conditions. This is because a multiplicity of contours exists in an image with different orientations and spatial scales. To research contour perception in natural images the process of contour segmentation has been used (Elder & Goldberg, 2002). This way, an observer traces the contours visible to them, painting a more complete picture of their perception. Thus, we will present observers with natural stimuli and will instruct them to segment the outlines of the image contents. For this, we have created a segmentation tool, described in the following section.

Figure 4: Comparison of edges and contours of a natural image from the database of Grigorescu et al. (2003). To use edges edges as defined by the Canny edge detector would not be a good way to probe human edge detection, as human observers could not segment an image this way.

2.1.3 Segmentation Tool

Our objective is to enable our observers to segment contours in natural images. To accomplish this, we have developed our own segmentation tool, even though several tools already exist. The rationale behind this decision was that the available tools tend to be too general and we wanted to streamline the workflow for our specific experimental design. While developing our segmentation tool, we have drawn inspiration from existing tools and conducted multiple trial-and-error iterations to arrive at the following specifications.

Firstly, the tool should display the stimulus and allow participants to add visible contours on top of it using the mouse. Secondly, participants should be able to click individual points to add contours instead of having to free-draw, as drawing freehand can be challenging. For this, we took inspiration from the segmented line tool in ImageJ[2]. Thirdly, the tool should have an "undo" function that allows participants to remove painted contours without restarting the entire trial. Finally, participants should be able to decide when they are finished, and the resulting contour image should be saved to disk.

The segmentation tool is operated using a mouse and a 5-button controller. The workflow for a single trial is as follows.

The stimulus is loaded from a mat file. The stimulus is accurately displayed at full size with its predefined properties. The remaining screen is grayed out. The gray level is the mean of the

[2] https://imagej.nih.gov/ij/docs/tools.html

displayed stimulus. Using the computer mouse the observer can mark the currently displayed stimulus via a segmented line tool. Segmented lines are started and extended using the left mouse button and ended with a right mouse click. The left button on the controller undoes the last mouse click, removing either a line segment or the endpoint of a segmented line. If the segmentation of one stimulus is finished, the current segmentation can be saved using the center button of the controller. The next stimulus is shown afterward. If all stimuli have been shown, the experiment exits.

The experiment outputs the saved segmentations and an output file. The segmentations are saved as binary images, with black being the markings and white being the background. As the contour maps from the natural image dataset are saved in the same manner, this allows for easy comparison. For each shown stimulus the output file contains: A trial number, the used image, the segmentation file, the mean and RMS-contrast of the stimulus, the RMS-contrast and mean luminance of the image, the noise type, and the RMS-contrast and mean luminance of the noise mask. An explanation of these values follows in section 2.2.2. Figure 5 shows a segmentation with the tool and its output binary image.

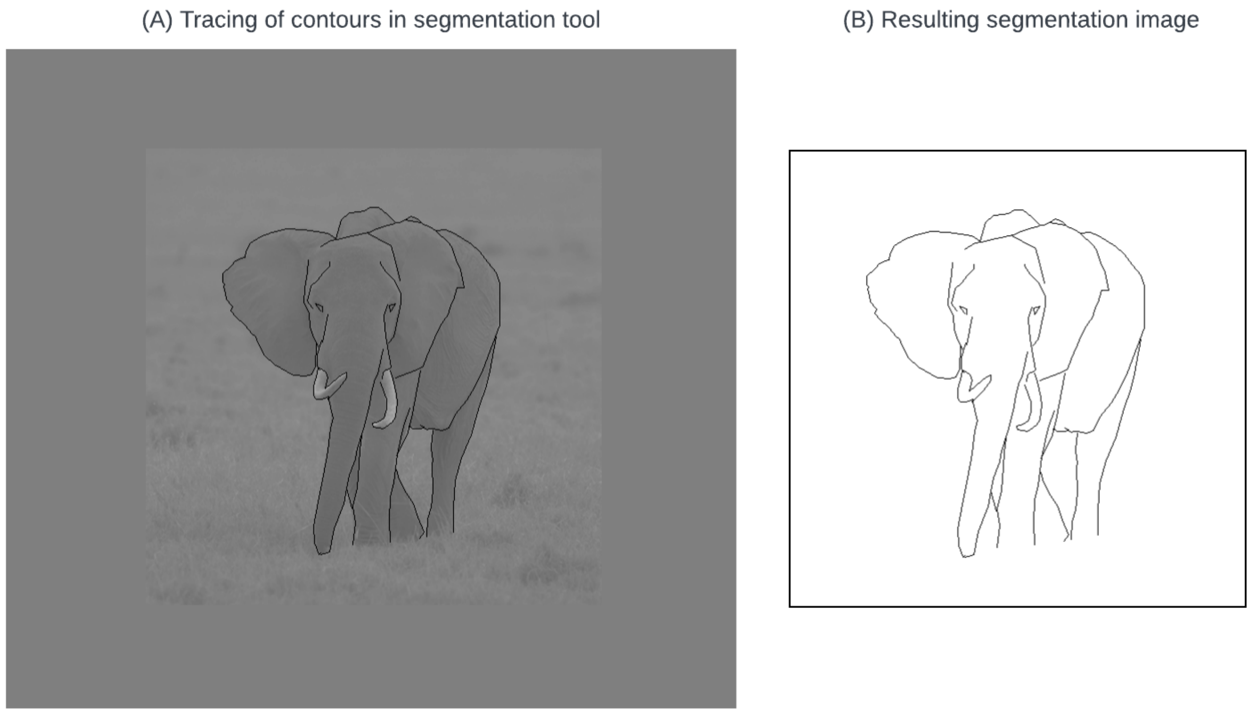

Figure 5: Example Segmentation. (A) shows the segmentation in the segmentation tool. (B) shows the resulting binary image.

2.1.4 In Search of a Segmentation Quality Measure

Now that we have a way of segmenting the contours in the natural images from Grigorescu et al. (2003), we need a way to measure the quality of a contour segmentation. We do this by comparing the segmentations with ground truth. Ground truth is here approximated by a human-labeled contour map on an undistorted image. Hence, a ground truth can either be the contour maps of the dataset by Grigorescu et al. (2003) or any segmentation of the same natural image, e.g. different trials of the same observer. However, for this, we further need a way to quantify the similarity between a segmentation and a ground truth.

One way to quantify this could be to correlate both images. The Pearson correlation coefficient is a measure of linear correlation between two random variables and can be used to measure how similar two signals are. It is defined as

$$\rho = \frac{cov(X,Y)}{\sigma_X \sigma_Y},$$

where X and Y are random variables, $cov()$ the covariance and σ the standard deviation.

If both images are flattened, the resulting vectors can be correlated using the Pearson correlation coefficient. This results in a value between -1 and 1, with 1 being a complete match between segmentation and ground truth and -1 for the segmentation being a negative of the ground truth. A blank segmentation would yield a result of 0. Since the task is to trace contours, which reflect the contents of the ground truth, a segmentation representing a negative of the ground truth cannot realistically be expected and hence results vary between 0 and 1.

The segmentations of observers are not pixel-perfect, therefore even if an observer segments the same edge as segmented in the ground truth, both will not necessarily align. This may pose a problem because it results in a lower score, even if the right edge was segmented. Potentially, producing inaccurate results.

Another approach to quantify the similarity between a segmentation and a ground truth is described by Grigorescu et al., 2003. They define the following similarity measure:

$$P = \frac{|E|}{|E|+|E_{FP}|+|E_{FN}|}$$

Where E refers to the set of correctly segmented contour pixels, E_{FP} to the set of contour pixels present in the segmentation but not in the ground truth, and E_{FN} to the ground truth pixels missed by the segmentation. We implemented this by looping through every pixel in the images and checking the conditions for each set to create them.

The advantage of this similarity measure is that we can set a margin of error. As it is very hard to do pixel-perfect, contour-aligned segmentations, an error margin has the benefit that results with small inaccuracies are still classified as correct segmentations. This way, a lower similarity score may not be caused by inaccuracies but by differences in perception. We implemented this as follows. First, we dilate both the ground truth and segmentation image by a factor. Thereby, the lines marking the edges will get thicker in both cases. When now comparing the not dilated images with the dilated ones, an error margin is introduced, as the segmentation in the not eroded image may now overlap with the eroded ground truth and vice versa. The error margin is the factor by which the images are dilated in pixels, e.g. with a factor of one, a segmented pixel in the segmentation is evaluated as correct even if it is one pixel away from a segmented pixel in the ground truth. From now on, we will refer to the process of comparing a segmentation to a ground truth using the similarity measure by Grigorescu et al. (2003) as the segmentation quality measure.

2.2 Test Case

After discussing some general considerations for an experimental paradigm on edge perception with natural images, we now explain in detail how we conducted our piloting procedure, addressing the efficacy of the experimental design and gathering preliminary data on edge perception in natural images. For this, we begin by describing the apparatus in section 2.2.1. Next, we describe the stimuli we used in section 2.2.2. Finally, explain our piloting procedure and the questions we wanted to answer in section 2.2.3.

2.2.1 Lab Environment / Apparatus

The apparatus consisted of a 21-inch Siemens SMM2106LS monitor (400 × 300 mm, 1,024 × 768 px, 130 Hz), a headrest, a common computer mouse, and a 5-button controller. The presentation was controlled with a DataPixx toolbox (Vpixx Technologies, Inc., Saint-Bruno, QC, Canada) and the presentation software HRL[3]. HRL is a library for running high-resolution

[3]https://github.com/computational-psychology/hrl

luminance experiments for psychophysics in Python. Unlike conventional computer monitors, the Siemens SMM2106LS monitor converts images from intensity values to luminance values linearly, i.e. the input values to the brightness values that the monitor physically emits. Additionally, the lab can be entirely darkened so that the only light comes from the monitor. This and the monitor both ensure that the stimuli are perceived at their intended luminance. A headrest at a distance of 100 cm from the monitor ensures the correct spatial frequency of the noises. The computer mouse and the 5-button controller (ResponsePixx button box, VPixxTechnologies, Inc.) are provided as input devices.

2.2.2 Stimuli

The goal of this research is to pilot an experiment that can be used to study edge perception in natural images and whether the findings from Schmittwilken and Maertens (2022b) and Shapley and Tolhurst (1973) about edge perception, that a narrow spatial scale around 3 cpd is most informative about edges, translates to natural images. To investigate the latter, we propose a noise-masking paradigm, i.e. we present the natural images masked with different kinds of noise.

For this, we use 6 different noise conditions and one control condition with no noise. We use three kinds of narrowband noise to trace which spatial frequency contents are most relevant for edge perception in natural images. These noises have center frequencies of 0.58, 3, and 9 cpd and a bandwidth of one octave, which is why we refer to them in the following as N0.58, N3 and N9. In addition, we use three broadband noises white, pink, and brown noise. We include those noises to test whether noises that cover a wide range of spatial frequencies affect edge perception in natural images in a different way, and hence to uncover potential nonlinear effects. Pink noise may be especially interesting as it models the frequency spectrum of natural images (Kayser et al., 2006). The basis for all noises is pseudo-random white noise with equal amplitude in all frequencies. For the narrowband noises, it is then band-pass filtered with a Gaussian filter with one octave spatial frequency bandwidth and respective center frequency. For the broadband noises, the amplitudes of the frequency components were divided by the frequency for pink (1/f) and by the square of the frequency for brown ($1/f^2$) noise.

All stimuli have a size of 12.8 x 12.8 deg at a viewing distance of one meter. Their mean luminance is 100 with the image having a mean luminance of 100 and the noise having a mean luminance of 0. The luminance range was between 22 and 182. We further define the RMScontrasts of the image and the noise. The RMS-contrast is the standard deviation divided by the mean. This enables us to control how visible the contours in the natural images are by changing their contrast (i.e. a higher image contrast would lead to more visible contours) or by changing the contrast of the noises (i.e. a higher noise contrast would lead to less visible contours). However, it is not clear how both the noise and image contrast should be chosen. Therefore, we first tested what a good value for the image contrast is. For this, we presented one randomly selected stimulus to one observer over a range of contrasts. The noise contrast was fixed at 0.15 and the image contrast was in a range between 0.01 and 0.09 with steps of 0.01 in between. The observer was instructed to segment the outlines of the image contents (2.1.2). The image contrast 0.01 should reflect floor performance (i.e. no visible contours) and 0.09 should reflect ceiling performance (i.e. noise does not hinder contour perception). For this piloting procedure, we used 63 stimuli (1 image · 7 noise conditions · 9 contrasts), which took the observer 2 sessions of approximately 2 hours to complete. Examples of these stimuli are listed in Figure 6.

2.2.3 Testing the Experimental Procedure

As previously mentioned, there are many things to consider if one wants to study vision with natural stimuli. We have developed an experimental paradigm with which contour perception in natural images can be measured and quantified. However, there were still some challenges and open questions that we needed to address before performing an actual experiment. These challenges summarized are, (1) whether the experiment works as intended, (2) what are appropriate contrasts, (3) and the relevance of the error margin in the measure of segmentation quality. To answer these we used the segmentations gathered in piloting (Section 2.2.2).
Due to the use of a new experimental paradigm, we are unable to predict its efficacy. To gain an idea of its potential, we compared the segmentations with the ground truths from Grigorescu et al. (2003) and the segmentations of the noiseless conditions as a ground truth. Our aim was to determine whether the scores of the segmentation quality measure match our perception of the similarity between two segmentations. Further, we wanted to test the similarity of segmentations from different trials. If the selection of contours varies too much between trials, differences in segmentation quality could solely be due to inconsistent selections. To examine this and to gain insight into the consistency between different trials, we first tested the similarity between the no-noise conditions of the segmentations qualitatively. Then, we

compared the similarity between two segmentations qualitatively and then quantified this with the segmentation quality measure. We compared the segmentations with the ground truth of Grigorescu et al. (2003) and with the segmentations of the no-noise condition with the same contrast. Next, we will address the open question of determining appropriate contrasts as already mentioned in section 2.2.2. As both the image and noise contrast influence the visibility of contours, they are important variables. To find good values for the image contrast, we here have fixed the noise contrast at 0.15 and varied the image contrast between 0.01 and 0.09 with steps of 0.01 in between.

In theory, contours do not have a thickness. However, if we extract contours by drawing them, these drawings will come with a thickness. This thickness could be chosen arbitrarily. If chosen very small, it is very unlikely that a segmentation will overlap with the ground truth because a pixel-perfect segmentation is hard to do for humans. In the segmentation quality measure, we take this into account through an error margin parameter. In a first attempt, we have chosen a value of 5 px as suggested by Grigorescu et al. (2003). However, we did not know whether this value also works with our segmentation data. Additionally, we wanted to see the influence of the exact parameter value on the final result. Hence we have varied it and looked at how the results change. We used three values for the error margin: (1) 0 px, (2) 5 px (Grigorescu et al., 2003), (3) 25 px (As 5 percent of the image size, is already pretty large). To further analyze whether qualitative differences exist, we computed Spearman's rank correlation coefficient between the results with the different error margin parameters. The Spearman rank correlation can be used to calculate the correlation between data of unknown distributions. A result of 1 would be a perfect positive correlation, -1 a perfect negative correlation, and 0 no correlation. Based on Zar (2005) the Spearman rank correlation can be computed as $\rho = \frac{6 \sum d_i^2}{n(n^2-1)}$, where $d_i = R(X_i) - R(Y_i)$ is the difference between the two ranks of each measurement and n is the number of measurements.

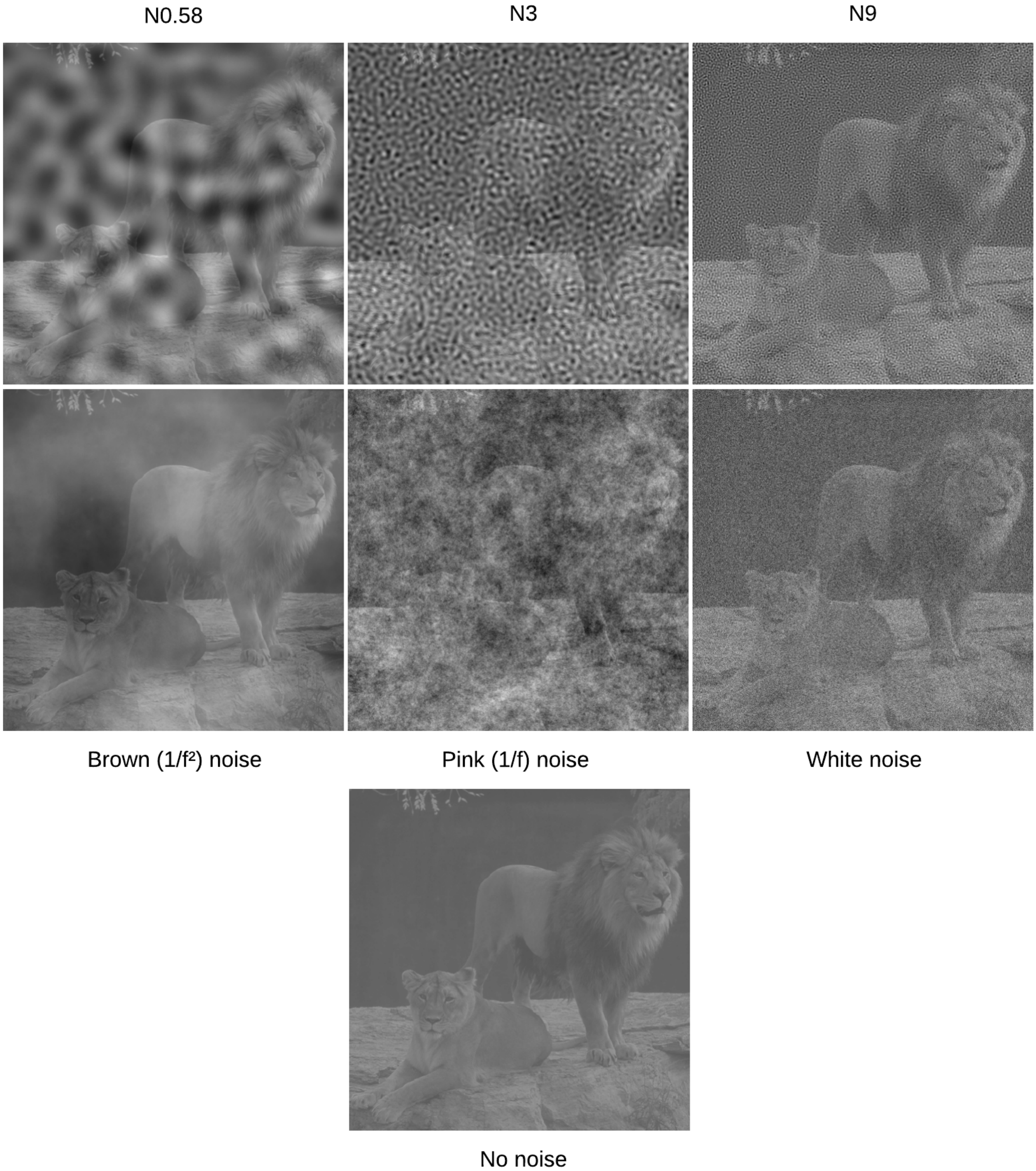

Figure 6: Examples of the stimuli used in the experiment. A natural image is masked with either N0.58, N3, N9, brown, pink or white noise. Additionally, a control condition with no-noise exists. In these demonstrations, we used the same image and noise contrast for all stimuli.

3 Results

3.1 Quantifying Segmentation Quality

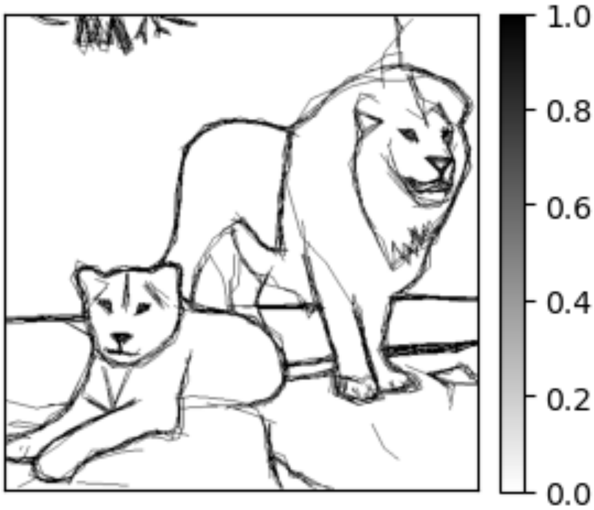

Figure 7: The segmentations of the no-noise conditions over all 9 contrasts layered on top of each other. The observer seems consistent in their selection of contours.

We have created an experimental paradigm to test human edge perception in natural images. To assess the efficacy of this paradigm we compare our piloting data qualitatively and then quantitatively using the segmentation quality measure. Figure 7 shows the segmentations of the 9 no-noise conditions from our piloting procedure layered on top of each other. Most lines have been drawn multiple times, with only a few single lines being visible. Hence, the observer seems to be consistent in their selection of contours between different trials. Figure 8 shows three example segmentations of noise conditions compared with the ground truths from Grigorescu et al. (2003) and no-noise conditions with the same contrast. Qualitatively, it is visible that the segmentations in the brown noise condition (Figure 8, rightmost image) are pretty similar to both the Grigorescu ground truth and the segmentation of the no-noise condition. Hence, we consider this to be a good segmentation which is also reflected in a high segmentation quality score (0.422 compared to Grigorescu ground truth and 0.498 compared to no-noise segmentation). Compared to that, we see that both the subjective quality of the segmentations as well as the segmentation quality scores are lower in the white noise condition (Figure 8, center image; scores 0.2 and 0.305), and worst in the N3 condition (Figure 8, leftmost image; scores 0.053 and 0.087). The scores resulting from a comparison with the Grigorescu ground truth are similar to the no-noise condition scores. They follow the same order as the no-noise condition scores but tend to be lower.

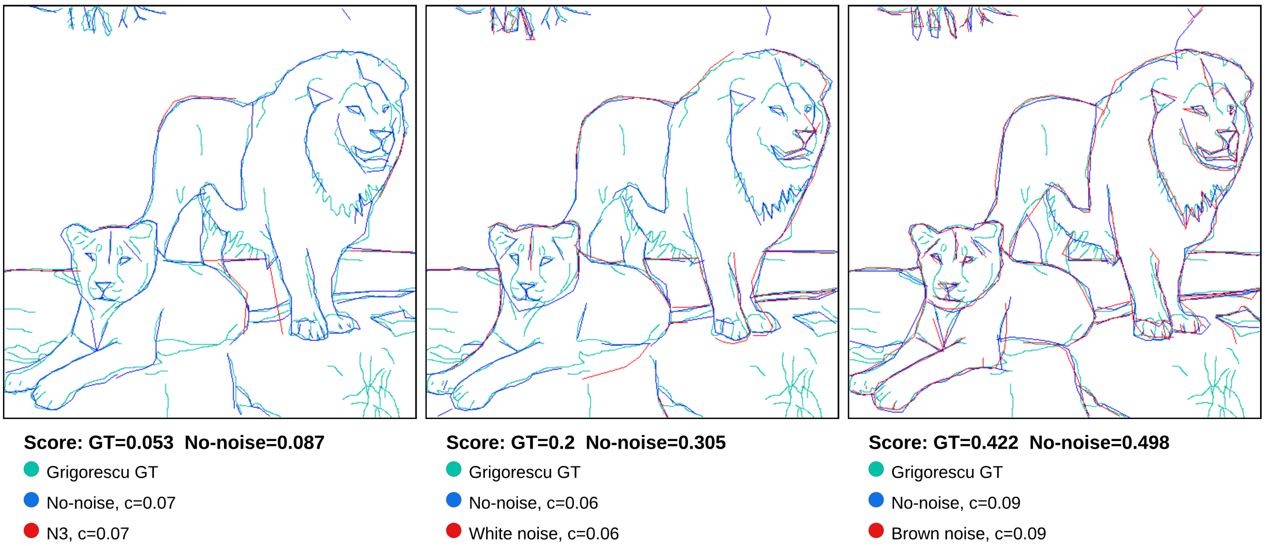

Figure 8: Comparison of noise condition segmentations (red) with Grigorescu ground truth (green) and the no-noise condition of the same contrast (blue). Below each image, the scores of the segmentation quality measure are listed. One for the comparison with the Grigorescu ground truth (GT) and one for comparison with the no-noise condition (no-noise). Below the scores, the contour maps are listed for each outline described by noise type and contrast or Grigorescu GT for the ground truths from the Grigorescu dataset.

3.2 The Effect of Image Contrast on Segmentation Quality

As a first step, we wanted to test the influence of the image contrast on the visibility of contours. Here, we quantified this by testing the effect of the image contrast on the segmentation quality. Further, we wanted to see whether the same effect as found by Schmittwilken and Maertens (2022b) translates to natural images, namely that pink noise and N3 noise interfere most with human edge perception. Figure 9 shows the segmentation quality scores over the contrasts for each noise type, using the Grigorescu et al. (2003) ground truth. Each line represents a specific noise condition. The x-value is the range of contrasts (0.01-0.09). The y-value is the result of comparing the stimulus with the contour maps as a ground truth using the segmentation quality measure. The segmentation quality of all noise conditions increases with higher contrast. The segmentation quality value of all conditions except pink and N3 shows a steep increase for contrast values between 0.01 and 0.06, thereafter increments are smaller. For pink and N3 no contours were segmented until a contrast of 0.04. Thereafter, segmentation quality increases slowly. Starting at a contrast of 0.02, a clear separation of the other noise conditions is visible. Pink noise has the overall lowest segmentation quality score closely followed by N3. The order thereafter is White, N9, N0.58, Brown and None.

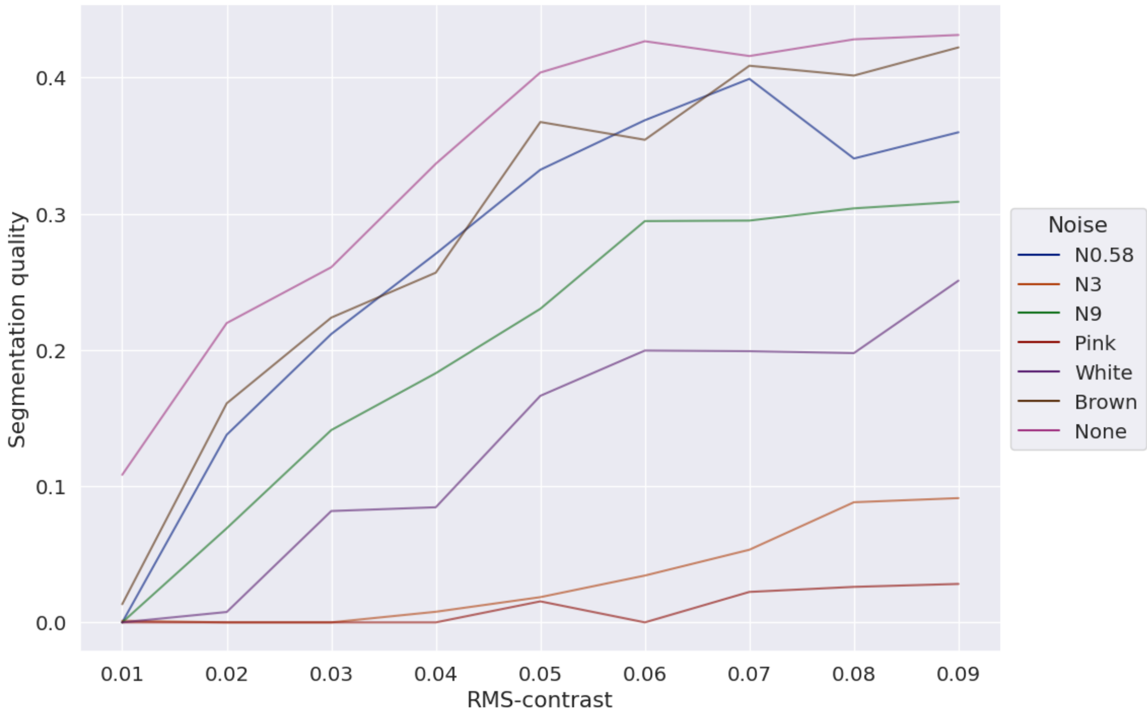

Figure 9: Results of the segmentation quality measure over the contrasts for each noise type, using the Grigorescu et al. (2003) ground truth and an error margin of 5 px.

3.3 The Effect of Different Error Margins on Segmentation Quality

In order to quantify the effect of the error margin on the segmentation quality score, we varied the error margin between 0, 5, and 25 px. Figure 10 shows a similar plot as Figure 9 but with a different error margin for each column. We find that the curves look visually similar between the different error margins. The noises are in the same order and the increase in the segmentation quality scores looks similar overall. However, the segmentation quality seems to correlate with the error margin resulting in higher segmentation quality scores with greater error margin values. For an error margin of 0, 5, and 25 the segmentation quality range is between 0 and 0.2, 0 and 0.4, and 0 and 0.8 respectively.

The results of the Spearman rank correlation are 0.995, 0.977, and 0.986 between error margins of 0 and 5, 0 and 25, and 5 and 25 respectively. Hence, the data across all error margins has a strong positive correlation.

Figure 10: Results of the segmentation quality measure over the contrasts for each noise type, with the Grigorescu et al. (2003) ground truth. Each column uses a different error margin.

4 Discussion

Edges are important features of our environment and are widely believed to be a first step in human visual processing. Evidence has accumulated that edge detection may be mediated by a narrow spatial scale around 3 cpd (Betz et al., 2015; Schmittwilken & Maertens, 2022b; Shapley & Tolhurst, 1973). However, this has only been confirmed using simplified stimuli. As these simplified stimuli do not necessarily unveil the inner workings of the human visual system (Olshausen & Field, 2005; Simoncelli, 2003; Touryan & Dan, 2001), complementary research using natural stimuli is necessary. However, while designing an experimental paradigm to investigate this question, we came to realize that testing human edge perception in natural images is a challenging task and that there is no standard approach for it yet. Hence, the aim of this thesis became two-fold. Firstly, we developed an approach to investigate human edge perception in natural images. For this, we proposed the task of contour segmentation. We quantified the quality of a contour segmentation by comparing it to a ground truth using a similarity heuristic specified by Grigorescu et al. (2003). Secondly, we wanted to investigate whether spatial frequency contents around 3 cpd also play an important role for human edge perception in natural images. For this, we used a similar noise-masking paradigm as Schmittwilken and Maertens (2022b), however, we applied this procedure to natural images. We then tested the experimental design using a noise-masked natural stimulus over a range of different contrasts, gathering data to answer the following: (1) How efficient is the experimental design? (2) What is an optimal noise contrast? (3) What is the influence of the error margin on the results?

We found that observers seem to be consistent in their selection of contours between different trials. The segmentation quality measure seems to match our perception of the quality of a segmentation. The scores of the segmentation quality measure are similar when comparing a segmentation to the segmentation of its no-noise condition and the ground truth maps from the dataset of Grigorescu et al. (2003).

The image contrast seems to have a strong influence on the results as the segmentation quality score correlates positively with the contrast. While we found that a wide range of image contrasts seem suitable for our specific noise-masking paradigm, contrasts below 0.04 should not be used because no contours could be segmented for the pink and 3 cpd noise conditions.

For the error margin, we found that the results for different values look qualitatively similar. This is supported by the almost perfectly positive Spearman rank correlation coefficients.

We will discuss these results and their implications on the experimental design in the following.

4.0.1 Validity of Experimental Paradigm

In this thesis, we have developed a method to test edge detection in natural images. The focus lies on contours which are defined as a subset of edges specifying the outlines of the image contents, not including edges originating from textured region. The reason for this is that it is not clear what an edge is in a natural image (Figure 4). Vision research on contours in natural images has been carried out using different approaches. Geisler and Perry (2009) used a contour occlusion principle, where parts of a contour were occluded and the task was to repaint them. As contours are occluded, this approach cannot be used to study contour detection. Another way to test contours is mentioned by Peters et al. (2005), they first show a natural image and then present an observer with an image of a segmented contour. The task is to judge whether the contour was present in the natural image. However, this may just give a small overview of contour perception as contours are diverse in their spatial scale and orientation and just one contour is tested for each image. We decided to segment contours in an image as employed by Elder and Goldberg (2002). Therefore, our experimental paradigm is based on the task of contour segmentation for research on edges in natural images. To measure the quality of a segmentation we decided to compare it with ground truth, as this seems to be a reliable way (if possible) to measure segmentation quality (Grigorescu et al., 2003; Huang & Dom, 1995). We approximate ground truth with human-labeled contour maps, which are either no-noise segmentations of the same observer or the ground truths by Grigorescu et al. (2003) The scores of the defined segmentation quality measure match our perception of the similarity between two contour maps (Figure 8) for both a comparison with Grigorescu ground truths and no-noise segmentations.

Comparing all no-noise segmentations, we found that the one observer that we tested seems to be consistent in their selection of contours (Figure 7). This consistency is coherent with the results by Martin et al. (2001) on closed contours. As all the data is gathered by one observer on one image, and the observer seems consistent on this one image this validates using no-noise segmentations as ground truth in this case. If the same holds true for multiple observers and different images, using no-noise segmentations of the same user as ground truth is a generally valid procedure because differences in segmentation can be mostly attributed to distortions of the image not inconsistencies in the selection of contours by one observer.

Since our data is currently limited to only a single observer, we could not quantify the similarity between different observers. As different observers could vary strongly in their selection of contours, this might be a limitation for using extraneous ground truths such as the Grigorescu et al. (2003) contour maps. Hence, it might have been favorable to analyze the contrast and error margin with no-noise conditions serving as ground truth. However, we compared the scores of the segmentation quality measure between different segmentations of noise conditions with either the no-noise condition or the Grigorescu contour map serving as a ground truth for three examples (Figure 8). In those cases both results are similar. For our procedure, this may be a qualitative argument for the use of the Grigorescu contour maps as ground truths. This is supported by Martin et al. (2001), as they measured the consistency of closed contour segmentations between different observers and found that they were highly consistent. However, to further validate the use of the Grigorescu ground truths or more generally extraneous contours, the interobserver variability would need to be measured.

Additionally, different noises and contrasts may influence other images differently. As has been shown natural image statistics are not entirely constant (Field & Brady, 1997), therefore using a range of different stimuli is important because results could be different for varying images. For these reasons, an experiment with multiple participants and a range of different stimuli would be vital. A range of stimuli would need to be selected. Especially, as different participants may judge contours differently. The results by Martin et al. (2001) indicate otherwise, but they segmented closed contours arguably leaving less room for error. Consequently, providing multiple contour maps per image to have a more representative human benchmark may be beneficial (Li et al., 2019; Martin et al., 2001).

Here we also want to mention that other experimental paradigms may have been possible with other advantages and disadvantages. For example, the spatial frequency of narrowband noise masks could have been varied by an observer until they perceived the strongest edge masking effect. This has the benefit that the spatial frequency response could be measured more accurately. However, it may also be complicated to test broadband effects with such a paradigm. To test edge perception in natural images further, it may generally be advantageous to employ a range of different methods. However, a definitive benefit of our approach is that we create contour maps of natural stimuli under different conditions. These could be used for further research, for example as a benchmark for physiologically inspired contour detection models (Grigorescu et al., 2003; Schmittwilken & Maertens, 2022a).

4.0.2 Choosing the Right Image Contrast

In our specific noise-masking paradigm, one open question was which contrast should be used, as the contrast influences the visibility of contours. If the contrast is too low no contours are visible, if the contrast is too high the contours are equally visible in every noise condition. Hence, it is important to find an appropriate contrast. As we define the contrast for the noise and image separately, in a first step we fixed the RMS-contrast of the noise at 0.15 cpd and the RMS-contrast of the image contrast varied between 0.01 and 0.09 with steps of 0.01 in between. Our results indicated that an image contrast below 0.04 should not be used with the noise masking paradigm we employed. In theory, one should also avoid that all contours are visible in all conditions. However, given the image contrast that we piloted, we found that we never reached ceiling performance (i.e. the noise is not hindering the segmentation), therefore a wide range of image contrasts seems possible to use. If one wants to plot psychometric curves (e.g. Figure 9) for all noise conditions, it would be desirable to have trials in which the observer does not see the contours in the image, and trials in which the observer sees all contours in the image and hence is at ceiling performance. Even though we tried a wide range of image contrasts, none of the image contrasts that we used was high enough to result in ceiling performance. However, in our specific case we could not increase the image contrast anymore, because that would have meant that some of the luminance values of our images would have been outside the range of luminance values that we could display on our monitor. One alternative that we did not test in our paradigm would be to reduce the noise contrast, rather than increase the image contrast. To test which noise contrasts could be used, the image contrast should be fixed at 0.09 and the noise contrast lowered stepwise, thus increasing the overall contrast.

4.0.3 The Relevance of an Error Margin

To test edge perception in natural images we proposed the task of contour segmentation, where an observer is instructed to segment the outlines of the image contents in an image. However, doing these segmentations it can be difficult to trace a contour completely pixel-perfect. Hence, a valid tracing is scored as wrong by the segmentation quality measure we used. This may not be desirable as a low segmentation quality score could be only due to inaccurate tracings and not less visible contours. Therefore, we take this into account with an error margin. Testing this parameter we wanted to find out, what error margin to use and whether the changes in scores are of qualitative nature. Comparing the results visually we found that in our test case, the different error margins did not change the results with respect to the question of edge detection performance. The noises were in the same order and the increase in the segmentation quality scores looked similar overall (Figure 10). The results of the Spearman rank correlation were 0.995, 0.977, and 0.986 between error margins of 0 and 5, 0 and 25, and 5 and 25 respectively, i.e. almost perfect. This further implies that no qualitative differences are introduced by the error margin. This means it is not necessary to use an error margin to be able to quantify the results with the segmentation quality measure. However, it can be used to scale the results. Hence, choosing an error margin that reflects the human perception of similarity the most could be sensible. For us, it was a margin of 5 px between the 3 tested margins in this case.

4.0.4 The Relevance of 3 Cpd for Edge Perception in Natural Scenes

Despite the small sample size, we attempted to draw preliminary conclusions on the overarching question behind this project, whether edge perception is mediated by a narrow spatial scale of 3 cpd. We have found that narrowband noise of 3 cpd and pink noise, were most obstructing to contour segmentation in our piloting procedure. We discuss the implications of this in the following.

The purpose and functionality of early human visual processing are a topic of frequent discussion. There is evidence that processing seems to rely on multiscale mechanisms (Elder & Sachs, 2004). In a multiscale system, the visual input is processed at different levels of detail (i.e.

spatial scales) as opposed to a single-scale system. Prior work implied that the visual system does not rely on the full range of spatial scales to solve one problem (Betz et al., 2015; Solomon & Pelli, 1994). This is important as a multiscale system with no subsequent scale-specific processing would not differ from a single-scale system (Schmittwilken & Maertens, 2022a). Edge detection as researched on simplified stimuli has been indicated to be mediated by a narrow spatial scale around 3 cpd (Betz et al., 2015; Schmittwilken & Maertens, 2022b; Shapley & Lennie, 1985). Our results support this notion as we showed that noise of a narrow spatial scale of 3 cpd effectively hinders contour segmentation in natural images in comparison to the other tested conditions, implying that image components of 3 cpd are of special importance for edge detection. As a narrow spatial scale seems important for edge detection, our results are compatible with the idea of a multiscale filtering model.

We also reproduce the result found by Schmittwilken and Maertens (2022b) that pink noise was very effective in masking edges. This is interesting as pink noise is not narrowband but broadband with a power spectrum of $1/f$. This implies that mechanisms are at work that respond to a larger bandwidth. Assuming, that the human visual system is multiscale with multiple frequency channels, there is evidence that the bandwidth of these channels might follow a $1/f$ distribution (O'Hare & Hibbard, 2011). This means that every channel is stimulated equally with a stimulus of a $1/f$ power spectrum (McDonald & Tadmor, 2006).

As contours are suppressed by pink noise inhibitive processes might be at work. McDonald and Tadmor (2006) implied that maximal suppression may occur when the population of neurons in V1 is stimulated equally. Hansen and Hess (2012) also suggested that noise with a frequency spectrum of $1/f$ might be obstructing to the detection of features in an image. Assuming a multiscale characteristic of the human visual system, they concluded that this effect may be explicable based on contrast gain control, where responses of one neuron are inhibited by a pool among different spatial frequency-sensitive channels. These pools dynamically adapt the neuron's contrast response function to the ambient contrast of the environment (McDonald & Tadmor, 2006). As pink ($1/f$) noise optimally stimulates these channels, the suppression is maximal. Other studies came to a similar conclusion (Haun & Essock, 2010; McDonald & Tadmor, 2006). However, this does not fully explain our results because in natural stimuli contours are already embedded in a $1/f$ environment. Since, natural images do not have the same contrast across the whole image, in a neighboring region of a contour the contrast may be lower. Indeed it has been shown that high-contrast patches in natural images are sparsely (non-Gaussian) distributed (Lee et al., 2003). The contrast across the noise masks seems more constant though, which makes sense as they are based on filtered Gaussian white noise. If the inhibitive neural pooling just happens across a certain neighboring region, this may possibly explain why pink (1/f) noise still has this masking effect in natural images.

4.0.5 Conclusion

Edges are vital visual features in our environment and are believed to play a significant role in human visual processing. Evidence suggests that edge detection is mediated by a narrow spatial scale around 3 cpd, but most research has been limited to simplified stimuli, i.e. isolated edges. To gain a deeper understanding of the visual system, we need complementary research using natural stimuli. However, as there is no standard approach for testing human edge perception in natural images, in this thesis we (1) developed and evaluated a method for testing human edge perception in natural scenes and (2) investigated whether human edge perception in natural scenes is similarly affected by noises that interfere with spatial frequency contents around 3 cpd as is the case for isolated edges. Our approach involves a contour segmentation task, where participants segment outlines of image contents in natural images using a self-created tool. To assess the quality of our method and human performance, we compared the resulting segmentation maps with a ground truth using a similarity heuristic. Assessing our method, first results indicated that our approach seems to be valid and effective to test edge perception in natural images. Consistent with prior literature, we found that also for more naturalistic stimuli, edge perception deteriorates most when interfering with image contents at a spatial scale of 3 cpd. However, data on the noise contrast is missing and the general limitations of a narrow test case apply.

The implications of the piloting data on visual processing complement prior literature. Hence using our experimental paradigm for research on human edge detection in natural stimuli could prove to be insightful.

References

Betz, T., Shapley, R., Wichmann, F. A., & Maertens, M. (2015). Noise masking of white's illusion exposes the weakness of current spatial filtering models of lightness perception. *Journal of vision*, *15*(14), 1–1. https://doi.org/10.1167/15.14.1

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological review*, *94*(2), 115. https://doi.org/10.1037/0033-295X.94.2.115

David, S. V., Vinje, W. E., & Gallant, J. L. (2004). Natural stimulus statistics alter the receptive field structure of v1 neurons. *Journal of Neuroscience*, *24*(31), 6991–7006. https://doi.org/10.1523/JNEUROSCI.1422-04.2004

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision research*, *22*(5), 545–559. https://doi.org/10.1016/0042-6989(82)90113-4

Elder, J. H., & Goldberg, R. M. (2002). Ecological statistics of gestalt laws for the perceptual organization of contours. *Journal of Vision*, *2*(4), 5–5. https://doi.org/10.1167/2.4.5

Elder, J. H., & Sachs, A. J. (2004). Psychophysical receptive fields of edge detection mechanisms. *Vision Research*, *44*(8), 795–813. https://doi.org/10.1016/j.visres.2003.11.021

Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. *Nature neuroscience*, *8*(12), 1643–1646. https://doi.org/10.1038/nn1608

Felsen, G., Touryan, J., Han, F., & Dan, Y. (2005). Cortical sensitivity to visual features in natural scenes. *PLoS biology*, *3*(10), e342. https://doi.org/10.1371/journal.pbio.0030342

Field, D. J., & Brady, N. (1997). Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research*, *37*(23), 3367–3383. https://doi.org/10.1016/S0042-6989(97)00181-8

Geisler, W. S., & Perry, J. S. (2009). Contour statistics in natural images: Grouping across occlusions. *Visual neuroscience*, *26*(1), 109–121. https://doi.org/10.1017/S0952523808080875

Grigorescu, C., Petkov, N., & Westenberg, M. A. (2003). Contour detection based on nonclassical receptive field inhibition. *IEEE Transactions on image processing*, *12*(7), 729–739. https://doi.org/10.1109/TIP.2003.814250

Hansen, B. C., & Hess, R. F. (2012). On the effectiveness of noise masks: Naturalistic vs. unnaturalistic image statistics. *Vision Research*, *60*, 101–113. https://doi.org/10.1016/j.visres.2012.03.017

Haun, A. M., & Essock, E. A. (2010). Contrast sensitivity for oriented patterns in 1/f noise: Contrast response and the horizontal effect. *Journal of vision*, *10*(10), 1–1. https://doi.org/10.1167/10.10.1

Heath, M., Sarkar, S., Sanocki, T., & Bowyer, K. (1998). Comparison of edge detectors: A methodology and initial study. *Computer vision and image understanding*, *69*(1), 38–54. https://doi.org/10.1006/cviu.1997.0587

Huang, Q., & Dom, B. (1995). Quantitative methods of evaluating image segmentation. *Proceedings., international conference on image processing*, *3*, 53–56. https://doi.org/10.1109/ICIP.1995.537578

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, *160*(1), 106. https://doi.org/10.1113/jphysiol.1962.sp006837

Kayser, C., Nielsen, K. J., & Logothetis, N. K. (2006). Fixations in natural scenes: Interaction of image structure and image content. *Vision research*, *46*(16), 2535–2545. https://doi.org/10.1016/j.visres.2006.02.003

Lee, A. B., Pedersen, K. S., & Mumford, D. (2003). The nonlinear statistics of high-contrast patches in natural images. *International Journal of Computer Vision*, *54*, 83–103. https://doi.org/10.1023/A:1023705401078

Li, M., Lin, Z., Mech, R., Yumer, E., & Ramanan, D. (2019). Photo-sketching: Inferring contour drawings from images. *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1403–1412. https://doi.org/10.48550/arXiv.1901.00542

Marr, D. (1976). Early processing of visual information. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *275*(942), 483–519. https://doi.org/10.1098/rstb.1976.0090

Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *Proc. 8th Int'l Conf. Computer Vision*, *2*, 416–423. https://doi.org/10.1109/ICCV.2001.937655

McDonald, J. S., & Tadmor, Y. (2006). The perceived contrast of texture patches embedded in natural images. *Vision Research*, *46*(19), 3098–3104. https://doi.org/10.1016/j.visres.2006.04.014

O'Hare, L., & Hibbard, P. B. (2011). Spatial frequency and visual discomfort. *Vision research*, *51*(15), 1767–1777. https://doi.org/10.1016/j.visres.2011.06.002

Olshausen, B. A., & Field, D. J. (2005). How close are we to understanding v1? *Neural computation*, *17*(8), 1665–1699. https://doi.org/10.1162/0899766054026639

Peters, R. J., Iyer, A., Itti, L., & Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision research*, *45*(18), 2397–2416. https://doi.org/10.1016/j.visres.2005.03.019

Salmela, V. R., & Laurinen, P. I. (2005). Spatial frequency tuning of brightness polarity identification. *JOSA A*, *22*(10), 2239–2245. https://doi.org/10.1364/JOSAA.22.002239

Schmittwilken, L., & Maertens, M. (2022a). Fixational eye movements enable robust edge detection. *Journal of Vision*, *22*(8), 5–5. https://doi.org/10.1167/jov.22.8.5

Schmittwilken, L., & Maertens, M. (2022b). Medium spatial frequencies mask edges most effectively. https://doi.org/10.13140/RG.2.2.11382.06726

Shapley, R., & Tolhurst, D. (1973). Edge detectors in human vision. *The Journal of physiology*, *229*(1), 165–183. https://doi.org/10.1113/jphysiol.1973.sp010133

Shapley, R., & Lennie, P. (1985). Spatial frequency analysis in the visual system. *Annual review of neuroscience*, *8*(1), 547–581. https://doi.org/10.1146/annurev.ne.08.030185.002555

Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Current opinion in neurobiology*, *13*(2), 144–149. https://doi.org/10.1016/s0959-4388(03)00047-3

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual review of neuroscience*, *24*(1), 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193

Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature*, *369*(6479), 395–397. https://doi.org/10.1038/369395a0

Tolhurst, D. (1972). On the possible existence of edge detector neurones in the human visual system. *Vision Research*, *12*(5), 797–IN1. https://doi.org/10.1016/0042-6989(72)90006-5

Touryan, J., & Dan, Y. (2001). Analysis of sensory coding with complex stimuli. *Current opinion in neurobiology*, *11*(4), 443–448. https://doi.org/10.1016/s0959-4388(00)00232-4

Zar, J. H. (2005). Spearman rank correlation. *Encyclopedia of biostatistics*, *7*. https://doi.org/https://doi.org/10.1002/0470011815.b2a15150