The processing of natural images in the visual system

OLGA DYAKOVA
Abstract

Any image can be described in terms of its statistics (i.e. quantitative parameters calculated from the image, for example RMS-contrast, the skewness of image brightness distribution, and slope constant of an average amplitude spectrum).

It was previously shown that insect and vertebrate visual systems are optimised to the statistics common among natural scenes. However, the exact mechanisms of this process are still unclear and need further investigation.

This thesis presents the results of examining links between some image statistics and visual responses in humans and hoverflies.

It was found that while image statistics do not play the main role when hoverflies (Eristalis tenax and Episyrphus balteatus) chose what flowers to feed on, there is a link between hoverfly (Episyrphus balteatus) active behaviours and image statistics. There is a significant difference in the slope constant of the average amplitude spectrum, RMS contrast and skewness of brightness distribution between photos of areas where hoverflies were hovering or flying. These photos were also used to create a prediction model of hoverfly behaviour. After model validation, it was concluded that photos of both the ground and the surround should be used for best prediction of behaviour. The best predictor was skewness of image brightness distribution.

By using a trackball setup, the optomotor response in walking hoverflies (Eristalis tenax) was found to be influenced by the slope constant of an average amplitude spectrum.

Intracellular recording showed that the higher-order neuron cSIFE (The centrifugal stationary inhibited flicker excited) in the hoverfly (Eristalis tenax) lobula plate was inhibited by a range of natural scenes and that this inhibition was strongest in a response to visual stimuli with the slope constant of an average amplitude spectrum of 1, which is the typical value for natural environments.

Based on the results of psychophysics study in human subjects it was found that sleep deprivation affects human perception of naturalistic slope constants differently for different image categories (“food” and “real world scenes”).

These results help provide a better understanding of the link between visual processes and the spatial statistics of natural scenes.

Keywords: natural scenes, image statistics, hoverflies, optomotor response, cSIFE neuron, sleep deprivation

Olga Dyakova, Department of Neuroscience, Physiology, Box 593, Uppsala University, SE-75123 Uppsala, Sweden.

© Olga Dyakova 2017

ISSN 1651-6206
ISBN 978-91-513-0032-0
urn:nbn:se:uu:diva-328041 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-328041)
To Karin Nordström and
Dan Larhammar
This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from the respective publishers.
Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH</td>
<td>Centrifugal horizontal</td>
</tr>
<tr>
<td>cSIFE</td>
<td>Centrifugal stationary inhibite</td>
</tr>
<tr>
<td>GLCM</td>
<td>Grey Level Co-occurrences matrix</td>
</tr>
<tr>
<td>H1</td>
<td>Horizontal sensitive 1</td>
</tr>
<tr>
<td>HS</td>
<td>Horizontal system</td>
</tr>
<tr>
<td>RMS</td>
<td>Root mean square</td>
</tr>
<tr>
<td>LPTC</td>
<td>Lobula plate tangential cell</td>
</tr>
<tr>
<td>R1-R6</td>
<td>Photoreceptor 1-6</td>
</tr>
<tr>
<td>VS</td>
<td>Vertical system</td>
</tr>
</tbody>
</table>
A glossary of terms

**Image statistics** - quantitative parameters obtained from an image.

**First-order image statistics** - quantitative parameters calculated from an image by using only values of pixel brightness regardless their position in space.

**Natural input** (or **natural images**) - images with a structure statistically similar to that we believe our visual system is adapted to.

**Natural scenes** (or **real-world scenes**) - images which represent natural environment (e.g. trees, bushes, fields, sky, sea etc.).

**Second-order statistics** - quantitative parameters capture the spatial relationships between the pixels.

**Slope constant of an average amplitude spectrum** (also called alpha-value or \( \alpha \)) - one of the second-order statistics calculated from an image amplitude spectrum after doing Fourier transform.

**Spatial frequency** \((f)\) - is a measure of how often sinusoidal components of an image repeat in space per unit of distance.

**\(1/f\) statistics** - a form of the average amplitude spectrum common for natural scenes.
Introduction

The world around us contains an enormous amount of visual information. All of this information is coded in such a meaningful way that our brains continuously obtain, select and minimise its redundancy to make behaviourally appropriate decisions (Attneave 1954, Barlow 1961, Field 1987). To understand these mechanisms better, it is important to find keys to decode the link between the environment and biological visual systems.

A possible way of doing this is to quantify the environment, define the natural input to visual systems and apply the proper experiments to examine how the visual system processes this information (Field 1989, Geisler 2008). Natural input (or natural images) can be considered as images with a structure statistically similar to that we believe our visual system is adapted to (Hyvärinen, Hurri et al.). To reduce possible confusion, let us call here images which represent nature (e.g. trees, bushes, fields, sky, sea etc.) as either “natural scenes” or “real-world scenes”.

Any image can be described in terms of its statistics, which are different parameters extracted from the image (Hyvärinen, Hurri et al., Field 1989, Pouli, Cunningham et al. 2011). Some image statistics, such as contrast or brightness, can be understood intuitively. Others, such as the slope constant of an average amplitude spectrum (often called the alpha-value ($\alpha$) which is a dimensionless quantity), are more complicated since they are obtained after an image transformation (e.g. Fourier transform). An average amplitude spectrum of a natural scene can be characterised by a function $1/f^\alpha$, where $f$ is spatial frequency and $\alpha$ is a slope constant, which was originally believed to be equal 1 for natural scenes (Field 1987, Tolhurst, Tadmor et al. 1992).

The visual cortex in mammals matches $1/f$ statistics, while their retinal ganglion cells work as the spatial filter which shows redundancy redaction (Atick and Redlich 1992, Barlow 2001, Simoncelli and Olshausen 2001). Fly retinas work in the same manner by acting as a filter for natural inputs by changing the incoming amplitude spectrum into a flat signal (i.e. $\alpha$ becomes 0). However, less is known about the coding of natural images by higher order neurons in insects. This question is addressed in this thesis by investigating the response of a novel higher order neuron (cSIFE, centrifugal stationary flicker excited) to natural images.

The fundamental principle of visual science is that visual systems are adapted through evolutionary and developmental processes to the statistical properties of the environments in which those visual systems must work.
(Barlow 1961, Simoncelli and Olshausen 2001). Analyses of many images from different categories have shown that the alpha-value of images belonging to the “natural scene” category is actually varied from 0.8 to 1.5 because these scenes are multifarious (e.g. they are sea, forests, fields, bushes, trees, sky etc.) (Tolhurst, Tadmor et al. 1992, Torralba and Oliva 2003). The mean α for natural scenes is close to 1.2 (Tolhurst, Tadmor et al. 1992, Torralba and Oliva 2003). Nowadays we humans live in buildings and observe anthropogenic scenes (e.g. cars) more often than natural scenes (i.e. we live in a very different environment than we did a thousand years ago.) However, psychophysics studies conclude that the output of the human visual system is tuned to alpha-value of natural scenes (Knill, Field et al. 1990, Tadmor and Tolhurst 1994). Moreover, images with natural α are perceived by human observers as more aesthetically pleasing and comfortable for viewing (Redies, Hänisch et al. 2007, Graham and Redies 2010, O’Hare and Hibbard 2013). Is there any condition which can change this preference?

Acute sleep deprivation is linked to vision and cognition (Silvia 2005, Killgore 2010, Kahn, Sheppes et al. 2013) (Harrison and Horne 2000, Jackson, Croft et al. 2008, Bixler 2009, Killgore 2010), therefore the hypothesis which is tested in this thesis is that sleep deprivation influences the perception of the alpha-value of visual inputs. Moreover, sleep deprivation is a challenge of modern life. Many professionals, such as journalists, rescue workers, drivers, nurses, pilots have to work during night shifts, so it is essential to investigate more deeply the relationship between sleep deprivation and vision to generate recommendations of how to improve quality of life, health and working performance for those who regularly experience sleep deprivation.

Our understanding of how visual inputs can affect behaviour in humans may also be applied to other animals such as hoverflies. These insects live in very cluttered environments, yet perform vital behavioural tasks such as inspecting flowers, searching for oviposition sites and defending territories at incredibly high speed (Van Veen and Moore 2004, Chandler 2010). For this reason, they are often used as models in vision research. However, the link between exact image statistics and behaviour in hoverflies is still unclear.

This thesis is therefore focused on gaining a deeper understanding of the link between natural images and visual responses in humans and flies. Although at the beginning the main focus planned to be on the slope constant of the average amplitude spectrum (α), other image statistics, such as RMS-contrast, skewness of image brightness distribution and some parameters related to image texture were also investigated. The results of this study, presenting in this thesis are very exciting, however, some of them are surprising and need further investigation.
Background

“Look deep into nature, and then you will understand everything better.”

Albert Einstein

Equations describing natural scenes

When we look at photographs of our world in all its beauty, we can detect and classify different scenes, name objects within them, identify their colours and even estimate an approximate distance between the objects and the camera (Torralba and Oliva 2003). Mathematical description of these processes is an entire world itself. Natural scenes are not random: they represent a certain structure with particular regularities which means they can be distinguished from, for example, man-made or artificial scenes (Ruderman and Bialek 1994). To understand the natural world surrounding us better, image statistics can be used (Elder, Victor et al. 2016). Indeed, photos of different environments (e.g. a forest or field) have quantifiable image statistics. These statistics are parameters that can be calculated from an image and which allow us to make inferences about a photo, to compare two or more images or to classify images into categories (van der Schaaf 1998, Pouli, Cunningham et al. 2011).

Image parameters can be classified by order (Field 1989, van der Schaaf and van Hateren 1996, Torralba and Oliva 2003, Pouli, Cunningham et al. 2011, Schwegmann, Lindemann et al. 2014). First-order statistics, such as RMS-contrast, skewness of brightness distribution, are aimed at examining the simplest image regularities and use the information about the brightness values of the individual pixels for analysis regardless of their position in space. Second-order statistics, such as the slope constant, or alpha-value, capture the spatial relationships between the pixels (Field 1989, van der Schaaf and van Hateren 1996, Pouli, Cunningham et al. 2011).

First-order statistics

Let us consider a simple greyscale image with only five levels of grey, where the pixels are randomly distributed in a 5x5 matrix (figure 1A), where the darkest (black) picture element, or pixel, has a value of 0 and the brightest (white) has a value of 255 (figure 1B). As can be seen in Figure 1B, each
pixel of this image has a specific luminance value that increases with the pixel's brightness. The resulting brightness distribution gives an overview of how many pixels with each intensity are present within the image.

**Figure 1.** A) An example of a simple greyscale image, with a size of 5x5 pixels. B) The luminance values of the image in panel (A) show that 255 corresponds to the brightest (white) pixel and 0 to the darkest (black) pixel.

To quantify the shape of the brightness distribution we use statistical moments (Pouli, Cunningham et al. 2010). The universal equation of any moment is:

\[ m_k = \frac{1}{N} \sum_{i=1}^{N} (x_i - c)^k, \]  \hspace{1cm} (1)

where \( k \) is the moment’s order, \( c \) is a constant, \( x_i \) is the intensity value of the pixel \( i \), and \( N \) is the total number of pixels (Pouli, Cunningham et al. 2011). Moments can further be divided into two groups: raw moments, where \( c=0 \), and central moments, where \( c \) is the mean brightness. The mean brightness is calculated from Equation 1 by using \( c=0 \) and \( k=1 \):

\[ \mu = \frac{1}{N} \sum_{i=1}^{N} x_i, \]  \hspace{1cm} (2)

The second central moment (\( c=\mu, k=2 \)) is variance.

The skewness (\( S \)) of the image brightness distribution is related to the third central moment (\( m_3 \)) of the brightness distribution (where \( k=3, c=\mu \)) and provides information about the relative amount of dark and bright pixels (Pouli, Cunningham et al. 2011):

\[ S = \frac{m_3}{\sigma^3}, \]  \hspace{1cm} (3)

where \( \sigma \) is the standard deviation.

Dark images have more positive skewness than bright images (Elder, Victor et al. 2016). The skewness of natural scenes is not symmetrical, and the
level of asymmetry varies between the studies. Natural scenes, particularly panoramic photos, are not symmetrical at all elevations (Schwegmann, Lindemann et al. 2014). As the sky dominates above the equator, this increases the number of brighter pixels, so the mean luminance will be higher in the upper part of the photograph (figure 2).

Figure 2. Example of variations of skewness and RMS-contrast at different elevation within one image.

The variation of brightness and contrast within the image also depends on the segment category, including backlit, sky, foliage, or ground (Frazor and Geisler 2006).

Contrast has many definitions in the literature (Barlow 1972, Bex and Makous 2002). The Weber and Michelson contrasts are often used to estimate the contrast of simple images (i.e. sinusoidal grating or patch of light on a uniform background) (Peli 1990).

The Weber contrast is a good metric for estimating the differences in contrast between the background and a feature (Peli 1990):

$$C = \frac{\Delta L}{L_{\text{background}}},$$

where $\Delta L$ is increment or decrement in the target luminance from the background luminance $L_{\text{background}}$.

The Michelson contrast is based only on the brightest and darkest pixels in the image, irrespective of the variation seen in between:

$$C = \frac{L_{\text{max}}-L_{\text{min}}}{L_{\text{max}}+L_{\text{min}}},$$

where $L_{\text{max}}$ is the maximum and $L_{\text{min}}$ is the minimum luminance in the grating.

Michelson contrast can measure the contrast of periodic images well (Peli 1990). However, because the Michelson contrast can be significantly increased or decreased by adding just one extremely dark or bright pixel in the image.
image, it does not provide good overall information regarding contrast in complex images, such as natural scenes (Peli 1990).

The most useful contrast for investigating natural images is the RM-contrast (Bex and Makous 2002), which is defined as:

$$RMS = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu)^2},$$

(6)

Just like skewness, the RMS-contrast of natural scenes depends on the elevation (Schwegmann, Lindemann et al. 2014).

Second-order statistics

Let us now compare two other images (figure 3A and figure 3B). By qualitative observation they are very different from each other, and the image in figure 3A can be perceived as more natural, while the image in figure 3B looks more artificial. However, by looking at the distribution of pixel brightness values and by calculating the skewness of the two images, we find that the two images are quantitatively the same (figure 4A and figure 4B; RMS contrast=0.22, skewness=1.17). What make the images in Figure 3A and Figure 3B look different to us? The image in Figure 3A has exactly the same pixels as the one in Figure 3B, however, the spatial location of the pixels in Figure 3B is randomized.

![Figure 3. A) An example of a natural scene. B) The same scene as in panel A, but after randomizing the spatial location of its pixels.](image-url)
**Figure 4.** Comparison of two images by their first-order statistics. A) The brightness distribution of the pixels shown in Figure 3A. B) The brightness distribution of the pixels shown in Figure 3B.

To take into account the spatial distribution of pixel brightness, second-order statistics, such as the amplitude spectrum of an image, are usually used (Field 1989, Pouli, Cunningham et al. 2011). To obtain such statistics, a Fourier transform is required. Any signal, including two-dimensional array of brightness values, \( f(x,y) \), such as an image, can be represented as a combination of a set of sinusoidal waves of different frequencies with varying amplitudes and phases (Gonzalez 1977). By using a Fourier transform any image can thus be represented in the frequency domain:

\[
F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-\frac{\pi i (ux + vy)}{N}},
\]  

(7)

where \( u \) and \( v \) are numbers of cycles fitting into one horizontal and vertical period of the image \( f(x,y) \). \( F(u,v) \) is the Fourier matrix consisting of complex numbers. \( N \) is number of pixels.
\[ A = \sqrt{R(u,v)^2 + I(u,v)^2} \]  

(8)

and

\[ Ph = \arctan \left( \frac{I(u,v)}{R(u,v)} \right) \]  

(9)

are the amplitude and phase of an image, where \( R(u,v) \) and \( I(u,v) \) are real and imaginary parts of complex numbers. The arrays of amplitude and phase together define the frequency spectrum of an image. The amplitude spectrum corresponds to sinewaves of the image and the phase spectrum to their relative shifts or orientations (Gonzalez 1977). Thus, by destroying (e.g. randomizing) the phase spectrum, an image will be unrecognizable (figure 5A). However, if the process is repeated with the amplitude spectrum we are still able to identify the scene within the image (figure 5B).

Figure 5. An examples of swapping phase and amplitude between two images. A) amplitude is taken from natural scene and phase is taken from random noise. B) phase is taken from natural scene and amplitude is taken from random noise.

After doing any kind of manipulation in the Fourier domain any image can be reconstructed from its frequency spectrum by using an inversed Fourier transform (Gonzalez 1977).
To quantify the amplitude spectrum with a single value, the slope constant ($\alpha$) of the average amplitude can be used. For this, the average amplitude spectrum across all orientations has to be calculated. A single value is useful for image comparison. For example, if the slope is steep, i.e. $\alpha$ is high, the image consists of less fine details (Torralba and Oliva 2003). When the average amplitude spectrum of natural images are plotted on a log-log scale there is a linear relationships between the amplitude and the spatial frequencies (Field and Brady 1997, Pouli, Cunningham et al. 2011):

$$A(f) = c/f^\alpha,$$

(10)

where $A$ is the amplitude, $f$ is the spatial frequency and $c$ is a constant.

Now, let us look at the average amplitude spectra of the images from Figure 3A (black data, figure 6) and from Figure 3B (grey data, figure 6). We can see that the amplitude spectra are different and that the amplitude of the original image (black data, figure 6), the one which qualitatively can be considered as more natural (figure 3A), is steeper (i.e. it has a higher slope constant).

![Figure 6. The averaged amplitude spectra across all orientations for the images in Figure 3A (black) and Figure 3B (grey).](image)

A previous study have shown that the slope constant of the amplitude spectrum typical for natural scenes is distributed between 0.8 and 1.5 with a peak of around 1-1.2 (Tolhurst, Tadmor et al. 1992). This variation can be explained by the wide variety of orientations and amplitudes in spatial frequency distributions in natural scene categories (e.g. fields, mountains, rivers, and natural objects such as flowers etc.) (Torralba and Oliva 2003). Torralba and Oliva (2003) described different scene categories and provided a good illustration of the amplitude spectrum. They suggested that distant scenes have a high impact of the sky and that makes them different from the close-up scenes,
which are isotropic in signatures in the spatial domain (Torralba and Oliva 2003).

Other image statistics that take the spatial distribution of the pixels within an image into account are parameters calculated from the Grey Level Co-occurrences matrix (GLCM) (Korchiyne, Farssi et al. 2014). This matrix represents how often pixels of particular values occur in the image. Once the matrix is created from an image it is possible to calculate parameters such as entropy, which denotes image randomness and complexity of image texture, energy, which is a measure of image constancy, correlation, which describes the consistency of image texture, and homogeneity, which shows local changes within the image (Selvarajah and Kodituwakku 2011, Zhao, Shi et al. 2014). A Grey Level Co-occurrences matrix has been applied for different classification tasks, including flower classification (Guru, Kumar et al. 2011).

Hoverflies and their behaviour

Hoverflies, including _Episyrphus balteatus_ and _Eristalis tenax_ (Diptera, Syrphidae) are commonly found in woods, gardens, near ponds, parks or in fields (Van Veen and Moore 2004). Hoverflies are so named because of their specific flight pattern in which they hover nearly motionless for prolonged periods of time (Fitzpatrick and Wellington 1983). However, they are also able to perform sidewise, backwards and turning movements (Collett and Land 1975).

Fly’s behaviour can be considered as unconstrained behaviour and behaviour under controlled conditions.

All examples of unconstrained behaviour of adult dipteran can be placed in four groups: reproduction (swarming, territorial, courtship), oviposition (insect host, plant host, animal host), survival (feeding, migration, mimicry, hibernation) and secondary effects of behaviour (disease vector and pollination) (Chandler 2010). The detailed behaviour classification is different for different species.

Fitzpatrick and Wellington (1983) focused mostly on the territorial behaviour of large hoverflies, including _Eristalis tenax_ (Fitzpatrick and Wellington 1983). They suggest that these hoverflies settle within an individual home-range, which can be defined as a large living area, including sites for different activities, such as resting, basking, grooming, feeding and territorial behaviour (Burt 1943, Wellington and Fitzpatrick 1981). According to their classification, territorial behaviour is divided into two groups: on-duty, when male hoverflies respond to intruders, and off-duty behaviours, when male hoverflies do not respond to intruders. Examples of on-duty behaviour are watching, inspecting and patrolling. Examples of off-duty behaviour are sitting or so called “go to” (Fitzpatrick and Wellington 1983). Watching and sitting behaviours
can be combined with basking, feeding or grooming. When a hoverfly is flying from point to point by seemingly random paths without responding to conspecifics they display “go to” behaviour. In contrast, inspecting and patrolling behaviours are characterized by the hoverfly following established paths (1 or more), either approaching to a within short distance (around 5 cm) of an intruder (without contact) and then returning back to the initial position (inspecting), or flying around in a territory (patrolling) (Fitzpatrick and Wellington 1983).

Alderman focuses on hovering behavior of *Episyrphus balteatus*. While some species of hoverflies orient themselves in a particular direction during hovering, *Episyrhus balteatus* turns through the entire 360° (Alderman 2010). Alderman (2010) considered hovering behaviour of *Episyrhus balteatus*, together with basking and conspecific competition as a part of swarming behaviour and is usually observed in sun shafts in between trees (Alderman 2010).

Thus, while being in their natural environment, insects including hoverflies, must recognize objects against the background in order to detect targets and orient themselves within their visual surroundings (Collett and King 1975). Flies are known to become fixated on vertical bars and high-contrast landmarks (Collett and Land 1975, Sareen, Wolf et al. 2011). They also respond to small objects such as other animals moving independently from the background or stable objects which appear to move on the fly’s retina because the fly is moving itself (Borst 2014). However, a real world consists of many different variables and this makes complicated to find a link between sensory stimuli and exact behaviour. Thus, the experiments in which variables can be under control are required (Chandler 2010).

Optomotor, escape, landing and fixation responses are commonly studied at the lab under controlled conditions (Borst 2014). During flight and while walking flies need to stabilise their movement by synchronization of their movement with the moving surroundings. This is called the optomotor response, which contributes to hovering and helps to stabilize locomotion (Collett and King 1975). An experimental setup can illustrate this optomotor behaviour. A tethered fly is placed to the centre of a moving drum with stripes on the internal walls. When the drum rotates in one direction the fly tries to follow the direction of moving pattern. The movement of the vertical body axis is called “yaw”, while movement of the transverse body axis is called “pitch” and movement of the longitudinal body axis is called “roll” (Blondeau and Heisenberg 1982). While rotational optic flow is independent of the distance between a fly and the moving surroundings, translational optic flow is dependent on the distance between an observer and an object and is therefore ideal for indicating landing and escape responses. By applying proper experiments it is possible to define neural control elements for these responses. (Borst 2014).
Hoverfly visual system

A sensory system is a part of the nervous system responsible for processing sensory information. Sensory systems consist of sense organs and their associated central processing areas. Sense organs are anatomical structures containing receptor cells and non-neural tissues. Sensory receptor cells perform a sensory transduction by using sensory receptor molecules (which are particularly sensitive to the appropriate sensory stimulus) to convert external stimulus energy into an internal electrical signal (receptor potential). These cells encode information from the surroundings and this information is transmitted to the central nervous system (Kandel, Schwartz et al. 2000).

Hoverflies are very good models to study vision: despite their small brain size, they efficiently process visual information from the surroundings and respond to it amazingly well and incredibly fast. However, it should be taken into account that in spite of similarity between flies and humans in the visual neural circuits (Sanes and Zipursky 2010), hoverfly eyes have a completely different structure (Land and Nilsson 2012).

Before light is converted into nerve impulses, it passes the optical apparatus. The quality of the optical system, the angular spacing of the receptors, and the diameter of the photoreceptors all influence the limitation of spatial vision (Land 1997, Warrant 2010). In contrast to humans and other vertebrates, flies have two types of eyes: the ocelli, which are sensitive to brightness, and two compound eyes, which are suited for spatial vision (Hengstenberg 1993, Borst and Haag 2002, Borst, Haag et al. 2010). Each compound eye consists of many thousands of ommatidia (Borst and Haag 2002, Borst, Haag et al. 2010). Each ommatidium contains 8 photoreceptors, which discriminate changes in luminance coming through the lens. The lenses in the fly eyes are very small and they provide a spatial resolution limited to about 1 degree of the visual field of view, which is very poor in comparison to human eyes (figure 7) (Straw, Warrant et al. 2006, Land and Nilsson 2012). Spatial vision in hoverflies is enabled by R1-R6 outer photoreceptors which send their axons to lamina for connection with large monopolar cells (LMCs) and amacrine cells (Borst, Haag et al. 2010, Borst 2014).

After being processed in the fly peripheral visual system, the visual input processes to the visual ganglia, which consist of three layers of neuropile: the lamina, the medulla and the lobula complex (Borst and Haag 2002, Borst, Haag et al. 2010). The lobula complex has two parts: the lobula and the lobula plate, where the visual interneurons called lobula plate tangential cells (LPTCs) can be found (Borst and Haag 2002, Borst, Haag et al. 2010). These cells respond to vertically or horizontally oriented motion. The vertical system (VS) cells, as their name suggests, are more sensitive to vertical motion, while the horizontal system (HS), centrifugal horizontal (CH) cells and H1 cells, are all sensitive to horizontal motion.
Figure 7. An example of viewing a part of the object by human observer (A) and after filtering the same image in accordance to the fly optics (B).

HS cells depolarize to the wide-field motion in their preferred direction and hyperpolarize to other directions of motion (Borst and Haag 2002, Borst, Haag et al. 2010). These cells are associated with the “yaw” optomotor response, which is self-rotation around the vertical body axis (Hausen and Wehrhahn 1983, Krapp, Hengstenberg et al. 1998, Haikala, Joesch et al. 2013).

It has been proposed that the HS is possibly influenced by a recently discovered higher-order visual neuron, the centrifugal stationary inhibited flicker excited (cSIFE) neuron, in the hoverfly (Eristalis tenax) lobula plate (de Haan, Lee et al. 2013). This neuron is inhibited by stationary patterns regardless of their orientation and is excited by non-directional motion, which is unusual for LPTCs (de Haan, Lee et al. 2013). These properties, however, are dependent on the pattern wavelength (de Haan, Lee et al. 2013).

Natural scenes and vision

It has been proposed that the visual system has evolved to efficiently process visual information received from the environment (Barlow 1961, Atick and Redlich 1992, Land and Nilsson 2012). The link between the first- and second-order statistics and, vertebrate and invertebrate visual systems, has been investigated to understand this process better. It has been shown that the asymmetry of ON-OFF ganglion cells in vertebrates matches the asymmetry in brightness distribution in natural scenes (Ratliff, Borghuis et al. 2010). Similar findings were discovered in the fly peripheral visual system: motion detectors
in *Drosophila melanogaster* are optimized to ON-OFF asymmetry of the visual world (Leonhardt, Ammer et al. 2016).

The peripheral and central visual systems of vertebrates are adapted to the second-order statistics of natural scenes (Barlow 1961, Atick and Redlich 1992). For example, in the periphery, the vertebrate retina is tuned to a natural $1/f$ spectrum and filters such inputs in such a way that the spectrum becomes flat (Atick and Redlich 1992). In the central visual system, cortical cells in the visual area V1 optimally encode scenes with the same $1/f$ characteristic (Field 1987, Field and Brady 1997). LMCs work as whitening filters, which conforms with the theory of maximizing information (Van Hateren 1992).

Recently by psychophysics study which was focused on the first-order image statistics was concluded that humans surprisingly prefer images with unnaturally low skewness (i.e. those which were manipulated so that the distribution of image brightness was symmetrical) (Graham, Schwarz et al. 2016). However, tuning of the human visual system to the second-order $1/f$ statistics has been shown during last few decades (Tadmor and Tolhurst 1994, Párraga, Troscianko et al. 2000).

It is possible to manipulate the slope constant of a given image (Tolhurst, Tadmor et al. 1992). When we increase the slope constant the image looks blurrier to a human observer, and when we decrease it, the image appears as it was drawn by pencil with more fine details (figure 8).

*Figure 8.* An example of the slope constant manipulation. A) Original image. B) Slope constants of original image (grey line), image with slope constant of 0.5 (dashed black), image with slope constant of 1 (dotted black line) and image with slope constant of 1.5 (solid black line).
This image manipulation was used to design an experiment which showed that the human visual system is best at discriminating changes in second-order statistics in images where the slope constant is close to 1 (i.e. typical of most naturalistic scenes) (Tadmor and Tolhurst 1994). It was also shown that the spatial statistics not only influence the perceived contrast of an embedded feature, but that the strongest contrast suppression took place when the background had a slope constant of 1 (McDonald and Tadmor 2006).

Image statistics relate to our aesthetic perception. When we look at paintings drawn by professional artists or at photographs of nature or human faces expressing different emotions, we can get different perceptions of pleasantness depending on that what we see. The aesthetics theory have been evolved during centuries and one of the recent is the efficient processing theory of aesthetics (Renoult 2016).

Recent studies have analysed the amplitude spectrum of art of different styles and epochs (Redies, Hänisch et al. 2007). They found that most art has similar regularities to images of natural scenes (Redies, Hasenstein et al. 2007, Graham and Redies 2010) (e.g. the slope constant of the amplitude spectrum varies but is close to 1). Moreover, by analysing paintings of human faces it has been shown that α of these paintings is unexpectedly similar to natural scenes and differs from alpha-values which are calculated from photos of real faces (Redies, Hasenstein et al. 2007, Graham and Redies 2010). Indeed, art aims to stimulate the visual system. It has been suggested that the visual system’s adaptation to natural scenes underlies creative art with similar second-order statistics (Redies, Hänisch et al. 2007, Graham and Redies 2010). Recently it has been shown that visual discomfort is linked to the increasing amplitude of an image in the Fourier domain (Fernandez and Wilkins 2008).

Sleep deprivation

By empirical experience, one might notice that after a night of staying awake there is a temporary discomfort, including a difficulty to focus on texts, and the surround appears blurrier (as if it’s α was increased). Our circadian rhythms suggest to sleep during night time and normal sleep duration is estimated as at least 7 hours (Durmer and Dinges 2005, Benedict, Brooks et al. 2012). However, the average sleep duration in the Western world has decreased dramatically over the last 50 years (Bixler 2009). The process of sleep has provoked interest since Ancient Greece, however, sleep deprivation started to be studied in 19th century and already the first results showed the importance of sleep (Finger 2001). Indeed, sleep deprivation affects health and different tasks performance.

Sleep deprivation is associated with slower processing of more-detailed visual information and reduction of behaviour performance, which was demonstrated during the study on professional drivers (Jackson, Croft et al.
2008). Moreover, sleep deprivation provoke difficulty in facial emotions recognition (Van Der Helm, Gujar et al. 2010).

There is also a link between sleep loss and obesity. In sleep deprived condition people increase their attention on food (Cedernaes, Brandell et al. 2014). fMRI study showed greater neural activity in areas associated with reward, motivation and decision-making as a response to food stimuli after sleep deprivation comparing to the normal sleep night (St-Onge, McReynolds et al. 2012).
Aims

This PhD project is focused on deeper investigation the link between image statistics and visual responses. Hoverflies and human subjects were used in this work to redress some gaps in our understanding of visual processing. More specifically my thesis aimed to:

1. Determine a link between image statistics and hoverfly behaviour, in both unconstrained and under controlled conditions (Paper I, Paper II and Paper III);
2. Investigate if there is an influence of 1/f image statistics on the inhibition properties of a higher-order visual neuron, cSIFE (Paper III);
3. Investigate if there a link between human perception of 1/f image statistics and sleep deprivation (Paper IV).
Summary of main findings

1. Image statistics is linked to hoverfly behaviour (Paper II and Paper III).

The RMS contrasts of photos of the ground related hovering or flying behaviour of *Episyrphus* hoverflies were significantly different ($p<0.0001$, unpaired nonparametric Mann-Whitney test with Bonferroni correction for multiple comparisons). The average RMS contrast in panoramic photos taken from the viewpoint of hovering or flying *Episyrphus* hoverflies were significantly different ($p<0.05$). The average slope constant ($\alpha$) of photos of the ground from the viewpoint of hovering or flying hoverflies was significantly different ($p<0.0001$). However, the average slope constants ($\alpha$) of panoramic photos were not significantly different ($p=0.54$). Image skewness of photos of the ground from the viewpoint of hovering or flying hoverflies was significantly different ($p<0.0001$). The image skewness of panoramic photos was significantly different ($p<0.0001$).

The predicting models based alpha-value, RMS contrast and skewness of photos of the ground and the surround demonstrated that skewness is the best predictor of active behaviour (i.e. hovering vs flying) of *Episyrphus* (figure 9A).

The behavioural optomotor response is depended on the slope constant, and it is strongest when $\alpha$ is close to 1 (figure 9B).

2. cSIFE inhibition is influenced by $1/f$ image statistics (Paper III)

The cSIFE neuron is inhibited by stationary images and its inhibition is maximal when the slope constant of the amplitude spectrum of presented stimuli is close to the mean in natural scenes (i.e. $\alpha=1$) (figure 10).

3. Perception of $1/f$ statistics if affected by sleep deprivation (Paper IV)

The chosen slope constants when viewing natural scenes were significantly higher after sleep deprivation versus uninterrupted sleep (figure 11A).

The alpha reliability (i.e. the ratio between the chosen slope constant and the original slope constant of the image) when viewing natural scenes was significantly higher after sleep deprivation than after uninterrupted sleep (figure 11B).
Figure 9. A) The relationship between the AUC and the probability of predicting the correct behaviour in the independent photos, B) The accumulated yaw optomotor response after 10 s stimulation with a natural (open symbols) or artificial (filled symbols) image manipulated to have different α values.

Figure 10. The inhibition (blue) generated by the stationary images after manipulation of the amplitude spectrum of original natural scene (A) and random noise (B). Manipulated images have α=0, 1 and 2. Spontaneous rate in grey.

Figure 11. A) The chosen α after sleep deprivation versus uninterrupted sleep (p=0.0002, Wilcoxon matched-pairs signed rank test). B) The alpha reliability after sleep deprivation than after uninterrupted sleep (p=0.0006, Wilcoxon matched-pairs signed rank test).
This project included many methods and involved several scientists. Therefore, here will be provided a brief summary of key methods I used in the studies and leaded to the main results (table 1).

Hoverflies *Eristalis tenax* (**Paper I** and **Paper III**) and *Episyrphus balteatus* (**Paper I** and **Paper II**) (figure 12A and figure 12B) were used in this study.

![Figure 12. A) Eristalis tenax B) Episyrphus balteatus](image)

**Figure 12. A) Eristalis tenax B) Episyrphus balteatus**

**Table 1. Methods.**

<table>
<thead>
<tr>
<th>Methods</th>
<th>Short description</th>
<th>Aim</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hoverfly behaviour</td>
<td>Visits were identified as either a landing or as an approach (i.e. flying towards the flower to within 5-10 cm). Hovering was defined as being near stationary for a minimum of 60 seconds. Flying was defined as moving from point to point, without returning to a given starting position.</td>
<td>I</td>
<td>I, II</td>
</tr>
<tr>
<td>Hoverfly identification</td>
<td>Some hoverflies were caught with a net for visual examination. In some cases, they were filmed while flying or hovering, or took photos for further identification.</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Artificial flowers</td>
<td>Model flower lures were created using paper, with the colour verified with a spectrophotometer.</td>
<td>I</td>
<td>I</td>
</tr>
</tbody>
</table>
OD did observations only together with KN in Uppsala, 2016

Odour blends were added to microcentrifuge tubes and placed in the centre of each flower or in the ground just underneath it. The negative control was made as a 5cm diameter black circle with no odour compounds. 8 artificial flowers were placed equidistantly in random order in 2-3 circles with a 90cm diameter. One of the circles was used to quantify the odour, abiotic and visual cues. The other 1-2 circles were used for behavioural observations.

**Trackball setup**

Wing-fixed, tethered *Eristalis tenax* were placed on the trackball, 8 cm in front of the CRT screen. During each trial a panorama (natural scene or random noise) image rotated at 110° / 1s for 10s. Between trials the screen was left at mid luminance for a minimum of 2s.

Large-field stimuli moving horizontally on screen elicited optomotor response in walking fly. Since the fly was tethered, it rotated a ball by legs. Two optical sensors extracted from high speed gaming mice provided information about the movement of a styrofoam ball (1.45g, 50mm diameter), which was placed in a cup supported with an air flow from beneath. From the rotation of the ball “yaw” response was calculated separately for each trial and then mean total yaw was found.

**Photographs**

Photographs related to hovering and flying behaviour

8 bit full-frame digital single-lens reflex Nikon D700 camera with a resolution of 4256 x 2832 pixels was used. The focus of the camera was manually controlled to avoid the influence of image blurriness which could affect the slope constant (alpha-value).

Photos of the ground were obtained approximately 1 meter above the ground corresponding to the location where the hoverfly was observed to be either hovering or flying. The size of these photos corresponded to approximately 1 x 1.5 meters (ca. 53 x 80 degrees of the visual field).

The panoramic photos were centred on the location where the hoverfly was originally observed to be hovering or flying. The camera was placed on a tripod with a panoramic head, ca 1m above the ground, using a level. 11-12 evenly spaced photos (2832 x 4256 pixels) were taken to get by merging them the full 360 deg. coverage.

Photographs related to visiting behaviour (OD participated only in designing methodology)

Photographs of each flower were obtained with a Sony DSC-HX1 with and without a 10x10 cm dull grey fabric collar around the flower. Corolla and inflorescence shape were manually scored, using terminology.
### Image analysis and manipulation

| Image analysis of photos of the ground | Each photo was converted to grayscale in and cropped to 2832 x 2832 pixel squares which were analysed. These squares corresponded to 53 x 53 degrees of the visual field of view. Each photo was linearly rescaled to cover the whole dynamic range from 0 to 255. Before calculating the RMS contrast and skewness of brightness distribution, images were first low-pass filtered with a cut-off frequency of 1cpd to take the hoverfly’s optics into account, which corresponded to 53cpi for the ground photos. | I | II |
| Image analysis of photos of the surround | Each panorama consisted of 11-12 photos. Each of these panoramic photo was analysed separately and then average of each parameter was found across the all panorama. Each photo was converted to grayscale and cropped to 2832 x 2832 pixel squares, linearly rescaled from 0 to 255 and then analysed. The panoramic segments corresponded to 70 x 70 degrees. Before calculating the RMS contrast and skewness of brightness distribution, images were first low-pass filtered with a cut-off frequency of 1cpd to take the hoverfly’s optics into account, which corresponded to 70cpi for the panoramic photos. | I | II |
| Slope constant calculation | First, images were converted to greyscale and then after doing Fourier transform the amplitude spectrum was extracted. Then, the average amplitude across all orientations as a function of spatial frequency was quantified. The slope constant of the amplitude spectrum (α) was identified by fitting a linear function to the average amplitude spectrum. The detailed explanation of calculation an average amplitude spectrum can be found in Paper III, Supplementary materials. | I, II | I, II, III |
| RMS contrast and skewness of brightness distribution calculation | Before calculated these parameters, images were filtered. In Paper II images were low-pass filtered with a cut-off frequency of 1cpd to take the hoverfly’s optics into account, which corresponded to 53cpi for the ground photos and 70cpi for the panoramic segments. In Paper III images were band-pass filtered between 0.06 and 1cpd to take the sensitivity of cSIFE into account. | I, II | II, III |
| Manipulation of the slope-constant | First image was converted to greyscale. Then a two-dimensional Fourier transform was performed and the amplitude spectrum was calculated. Then the Fourier-transformed image was divided by its amplitude spectrum to get a flat one, with an α of 0. By multiplying the result with the coefficient \((1 + k f^{\alpha})\) where k is a constant, any desired image α was gener- | I, III | III, IV |
By then doing an inverse Fourier transform and rescaling the image matrix from 0 to 255, the images were recreated with a different $\alpha$.

The detailed explanation of an average amplitude spectrum manipulation can be found in Paper III, Supplementary materials.

<table>
<thead>
<tr>
<th>Psychophysics study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participants</strong></td>
</tr>
<tr>
<td><strong>Sleep conditions</strong></td>
</tr>
<tr>
<td><strong>Original images</strong></td>
</tr>
</tbody>
</table>

III IV
resolution of 4256 x 2832 pixels. The focus of the camera was manually controlled to reduce artificial blurriness.

| Procedure | The participant was placed in a dark room in front of a linearized 13-inch screen with resolution of 2560 x 1600 (Retina MacBookPro) with a viewing distance of ~30 cm, with no head or eye fixation. The participant had unlimited time to observe each presented image, but in total the experimental sessions lasted no longer than 45 minutes. For each experiment (sleep deprived, normal sleep, morning and evening session) the presentation order of the images was randomly selected, and the starting \( \alpha \) of each image was randomly selected. The participant was instructed to press the left or right arrow on the keyboard until the image presented on the screen appeared as naturalistic as possible, e.g. a manipulated image of a tree should look like a real tree. The right arrow increased \( \alpha \) by 0.1, up to a maximum of 1.7, and the left arrow decreased the \( \alpha \) by 0.1, down to a minimum of 0.5. The minimum \( \alpha \) selected by any subject was 0.8 and the maximum was 1.7. This alpha was only selected in 2.2% of the shown experiments. Once the image was chosen, the participant was asked how pleasant he found the selected image on a scale from 0 to 100, where 0 is absolutely unpleasant and 100 is extremely pleasant. Then, the next randomly selected image appeared. | III | IV |
Results and discussion

Is there a link between image statistics and fly behaviour (paper I, Paper II, Paper III)?

Unconstrained behaviour (Paper I and Paper II)

Hoverflies live in a very cluttered environment but their behavioural responses are incredibly fast. These flies are efficient pollinators and one of the visual tasks they have to perform is flower recognition. What makes some flowers more attractive for hoverflies was investigated (Paper I). Together with abiotic and olfactory sensory cues, such visual cues as reflected light, shape, size of a flower, and image statistics of the photos of the flowers were analysed.

As image statistics slope-constant of amplitude spectrum and parameters of image texture (i.e. homogeneity, entropy, energy, contrast and correlation) were calculated. Even though image texture analysis was used to classify flowers (Guru, Kumar et al. 2011) and alpha-value is known to be linked to efficient coding of visual information in flies (Van Hateren 1992), these image statistics are not significant features of the flowers to be attractive for hoverflies, while a combination of multimodal factors, such as colour, shape, size and sent make a flower attractive for them (Paper I).

However, a link between slope constant of the average amplitude spectrum, RMS contrast, skewness of brightness distribution and two other types of hoverflies behaviour, hoverflies and flying, was found (see Paper II, figure 3).

Photos of the ground above which *Episyrphus* was observed as hovering or flying and panoramic photos of the corresponded surround were taken.

Further image analysis showed that photos of the ground above which *Episyrphus* was flying has lower RMS-contrast, slope-constant of the average amplitude spectrum and skewness of brightness distribution comparing to those above which *Episyrphus* was hovering. Interestingly, alpha-value for flying behaviour was lower (mean $\alpha=0.8$) than the mean alpha-value among natural scenes ($\alpha=1-1.2$) (Tolhurst, Tadmor et al. 1992). Another surprising finding was that skewness of photos of the ground related to flying behaviour was negative, which is unusual for the most of natural images (Richards 1982, Ruderman and Bialek 1994). This can be explained with the increasing number of bright pixels, since most of these photos were taken in the open field under direct sun which leads to increasing of the mean luminance. Since all parameters were calculated from the same photo, they depend one on another.
Indeed, a correlation between skewness and alpha-value and skewness and RMS-contrast was shown (see Paper II, figure 4C and figure 4D). This can suggest, that the low of ground photos related to flying behaviour is the consequence of low skewness.

In contrast to photos related to flying behaviour, images of the ground above which *Episyrphus* was observed as hovering, have positive skewness as it is common for natural scenes investigated in previous studies (Ruderman and Bialek 1994). Indeed, photos of the surround correspondent to the hovering behaviour consist more trees (see Paper II, figure 1B), and the surround is more symmetrical (see Paper II, figure 2). This observation is consistent with the Alderman (2010) studies, where it was shown that *Episyrphus* prefer to hover during sunny days in areas surrounded by trees. This surround creates a specific pattern on the ground. Thus, photos of the ground related to hovering behaviour consist of more dark pixels than photos related to flying behaviour and explain the difference in skewness.

The panoramic photos of the surround were also analysed. RMS-contrast of the surround was significantly higher for flying behaviour, while there were no differences in alpha-value. The mean of the surround was 1.1, which is a common alpha-value for natural scenes (Tolhurst, Tadmor et al. 1992).

Since surround affects the ground pattern, it is really hard to say, statistics of which photos are primary for behaviour: ground or surround.

Parameters of ground and surround photos were used to create logistic-regression models to predict hoverfly behaviour. These models were then validated on a completely different photo set. It was concluded that combination of skewness of the ground and the surround is the best predictor. However, the only skewness of the ground can discriminate amazingly well between two behaviours.

This results, at first sight, were really surprising. It is known, that human visual system is optimised to natural alpha-value (Párraga, Troscianko et al. 2000) and even artists create their work with $1/f$ image statistics (Graham and Redies 2010). However, real world has a much larger dynamic range of luminance than art works (Graham and Field 2007) and paintings show lower skewness than natural scenes (Graham and Field 2007), so in art works the relationship between alpha-value and contrast and skewness unlikely so direct as in real world (figure 4C and 4D in Paper II).

Thus, it is essential to quantify natural environment in relation to behaviour.

As a next step, it would be reasonable to calculate the parameters of image texture of photos from the dataset used in Paper II and calculate RMS-contrast and skewness of brightness distribution of photos which were used in Paper I. Then, compare these statistics (i.e. the slope-constant of an average amplitude spectrum, RMS-contrast) of photos corresponded to flying, hovering and visiting behaviour, including. Photos of the flowers used in Paper I
are close up photos. It is known, that the distance from a scene affects its slope constant of an average amplitude spectrum (Torralba and Oliva 2003). However, in our data set we have photos of the ground taken 1 meter above the same flowers, which were used in Paper I. So, it is possible not only investigate the link between image statistics and hoverfly behaviour in more details but also compare how RMS contrast, skewness and parameters of image texture are changed depending on the distance to a camera. Moreover, by comparing the statistics only of photos related to the visiting behaviour and taken from a different distance, it will be possible to track the constancy of the relation between image statistics and hoverfly choice.

To verify the findings, obtained in the field, it would be reasonable to perform in-door experiments to confirm what sensory cues induce hoverfly behaviour. For example, freely moving hoverfly can be placed in a virtual environment with controlled image statistics. Virtual reality is an innovative system, which is already used to understand fly behaviour. So, it will be possible to examine, if image statistics found during the field work could induce hoverfly behaviour (e.g. hovering vs flying).

Under controlled condition (Paper III)

A trackball experiment, where tethered hoverflies while walking on a ball, followed wide-field moving stimuli with different $\alpha$ was performed. By measuring the total yaw, it was found that the optomotor response in walking flies is influenced by second-order statistics and it is maximal when the slope constant of the amplitude spectrum of presented stimuli is close to the mean in natural scenes. i.e. $\alpha=1$.

The distribution of RMS contrast of 109 natural scenes was found. First, these images were pre-filtered to be relevant to walking behaviour. No correlation between optomotor response and RMS-contrast distribution, nor between optomotor response and RMS-contrast of the images used during trackball experiment was observed. Thus, it was concluded that it is alpha-value and not RMS-contrast influenced optomotor response in walking Eristalis.

It has been previously shown by psychophysics studies that output of the human visual system is tuned to $1/f$ image statistics (Knill, Field et al. 1990, Tadmor and Tolhurst 1994). Here it was shown that output of a hoverfly visual system is tuned to the same image statistics. Thus, these findings provide further similarity between human and fly visual systems.

Trackball setup can be used further to investigate the relation between optomotor response in hoverflies in image statistics in more details. During field study (Paper II) it was shown that skewness of image brightness distribution is the best predictor of hoverfly behaviour. Thus, by changing skewness of moving panoramas, it will be possible to see how this parameter influences the optomotor response in walking hoverfly.
Fixation response can also be investigated in a relation to image statistics. It is known that flies respond to objects that move against a background (Borst 2014). In the real world, these objects can be, for example, other moving animals or stationary objects such as a tree (or vertical bar in the study arena (Borst 2014)) which appears as moving on the fly’s retina thanks to its own motion. It was previously shown that the image statistics of the natural world have some variations (Torralba and Oliva 2003, Schwegmann, Lindemann et al. 2014). By changing image statistics of the background and using trackball setup it will be possible to define how the statistics of the background influence the fixation response of walking hoverfly. Moreover, the statistics of the object (e.g. vertical bar) can also be manipulated and discrimination thresholds between the bar and a background can be found for different image statistics similar to the previous studies in human subjects (McDonald and Tadmor 2006).

Is there a link between second-order image statistics and higher-visual processes in insects? (Paper III)

The inhibitory property of the cSIFE neuron was investigated (Paper III, figure 1b). For this, intracellular recording while showing a range of stationary images was performed (Paper III, figure 1c). The inhibition varied for different scenes (Paper III, figure 1e) and this variation was correlated with second-order image statistics (i.e. $\alpha$) (Paper III, figure 2c). It was hypothesised that the cSIFE neuron is tuned to naturalistic scene statistics. To test this hypothesis, the amplitude spectra of several natural scenes were manipulated and images with new alpha-values ($\alpha = 0$, $\alpha = 1$ and $\alpha = 2$) were created. Also, an artificial image, which was a Gaussian noise image, was generated and its amplitude spectrum was manipulated in the same manner as amplitude spectra of natural scenes. By performing electrophysiology again, the strongest inhibition of cSIFE neuron to stimuli with $\alpha=1$ for all images, including the artificial image, was observed (Paper III, figure 3). To investigate whether this inhibition was the consequences of a high image contrast the cSIFE response to manipulated images was plot as a function of RMS contrast. No correlation between the inhibition of cSIFE and the image RMS-contrast was shown (Paper III, figure 4a).

Natural scenes are not random and they consist of an enormous amount of visual inputs. To encode this efficiently, the visual system should either maximize the information in relation to noise, or, conversely, reduce its redundancy. However, these mechanisms complement each other (Barlow 1961, Van Hateren 1992, Field 1993, Barlow 2001).

cSIFE neuron has a limited bandwidth and tuned to natural $1/f$ statistics. This neuron receives input from the periphery (i.e. photoreceptors and LMCs),
which are adapted to $1/f$ image statistics and work as bandpass filters at high light levels (Van Hateren 1992, Van Hateren 1993, Van Hateren 1997). Thus, it can be suggested that they contribute to the $1/f$ tuning of cSIFE.

cSIFE is a novel neuron and its properties need to be more investigated. First of all, it is important to determine what exact processes contribute to the spatial frequency tuning of cSIFE. Second, the link between cSIFE inhibition and natural scene statistics can be investigated further by using stationary stimuli with given natural statistics obtained during field study (Paper II). Next, cSIFE neuron is excited by nondirectional motion (de Haan, Lee et al. 2013) and this property should also be investigated in relation to natural visual input. Finally, the role of cSIFE neuron in behaviour has been proposed (de Haan, Lee et al. 2013) and should be investigated. It can be done, for example, by combining trackball experiment and electrophysiology. Stimuli can be generated with the natural statistics related to either hovering or flying behaviour (Paper II) and intracellular recording can be performed in walking hoverfly.

Is there a link between second-order image statistics and sleep deprivation?

In the modern society, many people, especially hospital staff, pilots and drivers, require night shifts or long working hours. The awake night might lead to difficulty in focusing on texts and the surround appears blurrier. This could provoke not only temporal discomfort but also serious consequences, such as dangerous situations on the roads and traumas.

By psychophysics study the perception of natural scenes in humans under total sleep deprivation was shown. The slope constant of images was manipulated and it was quantified what slope constants human male subjects perceived as most natural after a night of sleep deprivation and compared this with their choice after a night with normal sleep. For this purpose, photographs of natural scenes and of food were used. Food images were used because sleep deprivation is linked to altered responses to food stimuli (Benedict, Brooks et al. 2012, St-Onge, McReynolds et al. 2012, Cedernaes, Brandell et al. 2014).

It was found that sleep deprivation affects the perception of naturalistic slope constants differently for photos of food compared with photos of natural scenes (Paper IV, figure 2). It was concluded that sleep loss does not affect the perception of slope constant of food as it is more relevant for survival under conditions of energy deprivation.

Furthermore, the human ability to discriminate the original slope constants of food images is lower after sleep deprivation (Paper IV, figure 4).
It was previously shown that scenes with natural second-order statistics appear as more aesthetically pleasant for human observers (Redies, Hänisch et al. 2007). To answer, whether sleep deprivation affects the aesthetical perception of alpha-value the participants were asked to rank images they selected as natural on a scale from 0 to 100 in terms of the pleasantness. Then pleasantness was plotted as a function of the selected slope constant. The results suggest that there are no differences in perception of pleasantness of alpha-value between sleep deprived and normal sleep condition neither for both scenes categories. However, images contained food were ranked as less pleasant than images represented the real world.

A recent study shows that image preferences related to skewness (Graham, Schwarz et al. 2016). In this study RMS-contrast was controlled but not the luminance. Moreover, the photos of food were taken from a different distance and angle than photos of natural scenes. These could explain the differences in perception between two image categories. However, it can be suggested that in sleep deprived condition aesthetic valuation of the stimuli with content related to the survival process not in a line with efficient processing theory of aesthetics.

Even though the results of the study need further evaluation, they are very exciting. Indeed, is a link between sleep deprivation and perception of images as natural was found.

It is very important to investigate deeper the relationship between sleep deprivation and vision to generate recommendations of how to improve life quality, health and working performance for those who constantly experience sleep deprivation. During the study, presented in this thesis, the greyscale images were used as stimuli. To make an experiment closer to reality and take into account that perception of food is linked to colour (Spence 2015) the psychophysics study should be repeated with colour images. Then, to find out whether the scene itself provokes the different perception of an image after sleep deprivation or the reason is in distance, viewing angle or other image parameters, including other image categories (e.g. man-made scenes, people, flowers) can help. Completely artificial images can be created with the parameters, which would be indicated as important and used as stimuli in another psychophysics test. Finally, using an eye-tracker can be reasonable to trace the eye movement across the scene to determine what part of the scene the participant is looking at while making the decision to calculate further image statistics of those parts of the image.
Conclusion

“The artist is a receptacle for emotions that come from all over the place: from the sky, from the earth, from a scrap of paper, from a passing shape, from a spider’s web.”

Pablo Picasso

When I started to work on this thesis in 2013, I planned to cover some gaps in our knowledge about visual responses and natural scene statistics. At that time, I could not even imagine how wide and vibrant this field is! Already our first results provoked new exciting questions and evoked even more curiosity.

I can’t stop saying to myself: “Look around! Our world is beautiful!” Indeed, our eyes are relaxed while looking at nature and stress dissolves when we are in a forest, near a lake or in the mountains. But what makes natural scene be aesthetically pleasant for us? What makes natural scenes look the way they do and be perceived differently from other scenes categories? How do the brains of different living organisms process all of this information? Are these processes similar for different creatures or dependent on their habitats and visual systems? Are these processes stable or can they be changed under certain conditions? I tried to address each of these questions during my PhD study.

In 2007 Redies proposed a model (Redies 2007), which predicts that human beings share a similar mechanism of aesthetic judgment. Specifically, he showed that aesthetical visual stimuli (i.e. art works) are coded more efficiently than artificial ones. His proposal was based on earlier studies showed that paintings created by artists have characteristics similar to those that nature has (Graham and Field 2007) and that neural networks resonate in response to natural scenes. Indeed, mechanisms of efficient coding have been studied in insects and vertebrates and the tuning of their visual systems to natural input has been demonstrated (Field 1987, Van Hateren 1992, Van Hateren 1992, Simoncelli and Olshausen 2001, Olshausen and Field 2004).

This thesis, “The processes of natural images in visual system”, provides additional insights into the link between natural inputs and the visual system in both humans and hoverflies. My hypothesis that a higher-order neuron in the hoverfly brain has a tuning to $1/f$ image statistics was confirmed and the study was amplified by showing that optomotor response in hoverflies is influenced by the same statistics. These findings show a salient likeness between higher order visual processing in insects and vertebrates. This complements
previous studies which showed similarity in their peripheral visual system (Tadmor and Tolhurst 2000, Clatworthy, Chirimuuta et al. 2003).

Surprisingly, field studies did not identify an exact link between $1/f$ statistics and hoverfly unconstrained behaviour. However, there is, indeed, a clear connection between image statistics and active behaviours in hoverflies. In particular, it is fascinating that by using only skewness of image brightness distribution, calculated from a photo of the ground, it is possible to predict very precisely, what behaviour we can expect of a hoverfly ($Episyrphus$), hovering or flying, above it.

Psychophysics studies showed that humans rate natural scenes as pleasant to view even when sleep deprived. However, the perception of scenes as natural is affected by sleep deprivation. This needs further investigation, because if natural scenes are perceived as blurrier after sleep deprivation this may have dangerous consequences for instance for tired drivers driving in fog or rain.
Acknowledgments

There are so many people I wish to say “thank you!” Without you, this would never be possible!

Karin Nordström and Dan Larhammar, my brilliant supervisors, for all of your help to develop my skills as a scientist; for being understandable and critical; for all of your support and patience. Karin, thank you for the given opportunity to work on this project and for giving me freedom and great guidance and encouragement at the same time in developing ideas! Malin Thyselius, Kristian Johansson and Josefin Dahlbom, my lovely lab members, for all of your help and company. Malin, for the great talks with you about hoverflies. Kristian, for the help with trackball setup. Christian Benedict, Frida Rångtell and Xiao Tan, for an amazing opportunity to do psychophysics study on sleep deprivation and vision. Thank you for all of your help, support and ideas. You are the best! Finn Hallböök, Helgi Schiöth, Cecilia Edling, Eva-Lena Wahlberg, Karin Nygren, Neil Ormerod, Berit Hård-Wallenquist, for your help and support with all of the administrative and paper work. Olof Nylander for the opportunity to teach. I loved it a lot! Kate Shebanits, Bo Xu, Helen Haines, Jan Hoeber, Anna Rostedt Punga, Minas Hellsand, Arianna Cocco, Laura Vossen, Olga Titova, Sveta Vasilovska for being around friendly and supportive. Yurii Bogatcev and Valerij Kiselev for the support in the first steps in science; Igor Kagan and Melanie Wilke for the support in the first steps in neuroscience. Victor Tuzov for the encouragement to the philosophical and psychological talks and thinking. Shannon Ohlson, Martin Egelhaaf, Martin Müller, Olivier Bertrand for the great collaboration on the fieldwork study. Adan Ulises Vargas and Malte Koester for a great time of working together and a friendship after that time has gone. Sam Hardman for help with English and enthusiasm. To all scientists, I mentioned in this thesis for the inspiration. Vova, for the help in a field and support in starting life in Sweden. Natasha Sugonyako, Olga Belova, Tomat Alexandra and Yulya for being kind, supportive and patient, when I almost gave up, you are the best girls, love you, really. Sashka for the love of coffee with milk, I was drinking it a lot while writing this thesis. My parents, for letting me to see this world. The world, for letting me be here and investigate it.

I might forget to mention many of wonderful people. However, if you read this, you should know that I am saying “thank you” to you too!


Field, D. J. (1989). What the statistics of natural images tell us about visual coding. *SPIE.*


faces with the Fourier statistics of complex natural scenes." Network:
Richards, W. A. (1982). "Lightness scale from image intensity distributions." Appl
Schwegmann, A., J. P. Lindemann and M. Egelhaaf (2014). "Temporal statistics of
natural image sequences generated by movements with insect flight
characteristics." PloS one 9(10): e110386.
features for content based image retrieval." International Journal of Latest Trends
in Computing 2(1).
appraisal theory of aesthetic emotions." Empirical studies of the arts 23(2): 119-
133.
(2012). "Sleep restriction leads to increased activation of brain regions sensitive
hoverfly (Eristalis tenax) eyes and associated faster motion detection and
Tadmor, Y. and D. Tolhurst (1994). "Discrimination of changes in the second-order
Tadmor, Y. and D. J. Tolhurst (2000). "Calculating the contrasts that retinal ganglion
cells and LGN neurones encounter in natural scenes." Vision Res 40(22): 3145-
3157.
Van Der Helm, E., N. Gujar and M. P. Walker (2010). "Sleep deprivation impairs the
van der Schaaf, A. (1998). Natural image statistics and visual processing, [University
Library Groningen][Host].
van der Schaaf, A. and J. H. van Hateren (1996). "Modelling the power spectra of
360(6399): 68-70.
A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)