Some Perspectives on Brightness Perception Models through the Study of Brightness Illusions

A thesis submitted to Indian Statistical Institute
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Computer Science

By

Ashish Bakshi

Under the Supervision of Dr. Kuntal Ghosh, Associate Professor

Machine Intelligence Unit
Indian Statistical Institute
Kolkata - 700 108, India

November, 2021

Some Perspectives on Brightness Perception Models through the Study of Brightness Illusions

A thesis submitted to Indian Statistical Institute
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Computer Science

By

Ashish Bakshi

Under the Supervision of Dr. Kuntal Ghosh, Associate Professor

Machine Intelligence Unit
Indian Statistical Institute
Kolkata - 700 108, India

November, 2021

ACKNOWLEDGEMENTS

This thesis is the end of my journey in obtaining Ph.D. This thesis has been seen through to completion with the support and encouragement of numerous people including my colleagues, well wishers, friends and relatives, and all those people who made the thesis possible. A few words cannot express the immeasurable value of the support I received. Thanksgiving is just a formality, and I do not know how to express my gratitude to all my beloved well wishers.

I would like to thank my supervisor Dr. Kuntal Ghosh who showed me how to navigate through the vast ocean that is scientific research. I would also like to express my immense gratitude to Dr. Alan Robinson of University of California, San Diego, for providing his codes for our use without any hesitation.

I would like to express my sincere gratitude to the late Prof. C.A. Murthy, without attending whose excellent lectures on Pattern Recognition and Image Processing I would have been unable to get a mathematical understanding and form a clear picture of the tools and methods that I have used in my research work. His insistence on mathematical rigor and completeness of knowledge has taught me what a scientist should strive for in research. Apart from these, his honest and upright manners, his student friendliness and his all time willingness to help out, has taught me how an ideal academic should conduct himself. His sudden demise has left a deep sense of loss in me.

I want to express my thankfulness to all the faculty members of my department. I would also like to acknowledge the support that I have received throughout from the office staff of our department during the tenure of my PhD. I express my sincere thanks to the authorities of ISI for the facilities extended to carry out research work and for providing me every support during this tenure. I want to thank all the others, whom I might have missed here, for their well wishes and support.

I want to thank all the teachers of my undergraduate and graduate studies, without whose guidance I might not have reached this position. I would like to specially express

my gratitude to the late Prof. Binayak Dutta Roy who taught me that there lies a simple and insightful understanding behind every deep concept in science. I would also like to specially mention the name of Prof. Albert C. Gomes and Prof. Ananda Dasgupta from my undergraduate years at St. Xaviers College who were the first teachers that I have personally known who demonstrated that it was even possible to have a deep and true understanding of the nature of the world around us.

I wholeheartedly thank all my friends for their constant encouragement and support, specially Ajoy Mondal, Bibek Sharma, Dhiman Karmakar, Jaydeb Mondal, Manish Chowdhury, Partha Pratim Kundu, Payel Sadhukhan, Ranajit Das, Sudeb Das, Sujoy Madhab Roy, Swarup Chattopadhyay, Tanujit Chakraborty and Tanmay Basu.

Ashish Bakshi

Abstract

Modeling brightness perception has always been a challenging issue in computer vision for performing real-life, complex visual tasks like object detection, recognition and image analysis under widely varying lighting conditions. Brightness constancy, brightness-contrast, brightness assimilation, transparency etc. are some of the crucial aspects in perceiving brightness that need to be dealt with in order to develop meaningful computational models in machine vision. The fascinating world of brightness illusions provides an important gateway to this study and research. The present thesis attempts to make some humble contributions to this end through both experimental psychophysics as well as from the perspective of computational neuroscience. In this work, we have, first of all, been able to design and establish four new illusory stimuli that expose the limitations of traditional spatial filtering models of vision in general and that of the well-established Oriented Difference of Gaussian (ODOG) filter, in particular. These four new stimuli are all modifications of some of the classical brightness perception illusions. The first is the Mach Band illusion where in our new design, the region of intensity gradient in the image increases linearly in size from zero at the top to its widest at the bottom, so that the bright Mach band is perceived thinner at the top and wider at the bottom diverging outwards like rays of light emanating from the top. Next, is the Hermann grid illusion where we introduced tiny perturbing squares overlapping with each grid square corner, resulting in complete wiping out of the illusory spots in Hermann grid. The third, deals with the sine and square grating stimuli that induce brightness to a foreground uniform test strip. Our design comprises a set of stimuli demonstrating sine to square grating transition unfurling new intriguing problems regarding the mechanism of brightness-contrast phenomenon in the parallel visual pathway and suggest that these two apparently similar effects of brightness induction may occur due to distinctly different mechanisms. The last one is the White's illusion where a simple longitudinal extension of the test patch in our modification demonstrates interesting effects in brightness assimilation and the failure of the ODOG model. We have shown that this

limitation of ODOG can be overcome by a new parsimonious spatial filtering model (Difference of Difference of Gaussian or DDOG) that can provide a unified explanation to both brightness contrast and brightness assimilation phenomena, and can also predict many such subtle brightness effects. Two versions of this model viz. the Varying Contrastive Context Filter (VCCF), and the Adaptive Contrastive Context Filter (ACCF) have been proposed, both of which attempt to approximate the Magno (M) and Parvo (P) channels in the central visual pathway. Next, some image processing applications of this new computational model in the domain of denoising with edge preservation are demonstrated especially in comparison to the powerful Bilateral filter. Finally, the limitations of our model and spatial filtering based approaches, in general, have been elucidated and the possible directions of future research in computational theories of brightness perception have been indicated.

Contents

List of Figures

List of Tables

Chapter 1

Brightness illusions and brightness perception

1.1 Motivation

Millions of years of evolution has equipped us with our five senses which we use to perceive the world around us. We usually trust our senses completely and we believe that whatever we perceive through the senses is reality itself. This belief is so strong that we even entrust our lives to it. For example, while crossing a road we rely on our senses to report the presence of an approaching vehicle. Reality however is far too vast and far too detailed to be perceived by our limited senses. Our eyes cannot sense any radiation beyond the visible wavelengths, which is only a thin slice of the electromagnetic spectrum. We also cannot perceive the fine details of the microscopic world which our unassisted eyes do not have the resolution to observe. We are therefore unable to perceive a majority of the signals that are impinging upon us because our sensory organs are ill equipped to sense them. Evolution has only equipped us with barely enough for our species to survive. From an evolutionary perspective dedicating excessive amounts of energy and nutrients towards maintaining sensory organs of excessively high power and acuity would be a waste of resources. Evolution has however equipped us with a brain whose job is to integrate all the incoming sensory information, draw inferences about the outside world from them, and decide upon the appropriate action depending on targeted goals. These inferences are nothing but estimates about the real world "out there" with whatever limited data we may be receiving from our senses. These estimates are what constitute our perception.

In most circumstances, our perceptual experiences are completely determined by the state of the external world that the brain intends to know about. Therefore, if we were to somehow record a person's perceptual experiences, then studying that record would only reveal information about the state of the external world and it will not reveal anything about that person's perceptual apparatus viz. the brain. Sometimes however, under certain circumstances, it has been found that the inferences drawn by the brain can be incorrect or inaccurate. These inaccuracies are anomalies in our perceptual experiences. Recording these anomalies would therefore reveal interesting information about the perceptual apparatus, instead of merely reflecting the state of reality. Sometimes these inferential errors are only temporary, but at other times they are persistent over time and consistent from person to person. These errors of estimation are termed as illusions. When they pertain to the sense of vision they are called visual illusions. These anomalies, i.e. illusions, offer us a window into the inner workings of the brain without having to physically probe it using instruments. Sometimes these perceptual anomalies also show a degree of variation from person to person. This variation too, may occasionally reveal information about the state of the perceptual apparatus of a person (e.g. detecting conditions such as colour-blindness, or even some diseases like diabetes [Davies and Morland, 2002]). By performing simple psychophysical experiments, we can gain insights into some of the underlying mechanisms of the brain. Psychophysical experimentation involves subjecting volunteers to a variety of external stimuli and recording their responses describing their subjective experiences of those stimuli. Insights into the human visual system and its underlying mechanism can also help in the development of new computational models of perception which can be important tools in solving various problems of computer vision and image processing. These

computational models can also assist us in mimicking the high robustness and sensitivity characteristics (such as 'identification', 'discrimination', etc.) of human vision.

Visual illusions have also found practical applications in various technologies. For example the appearance of motion in cinema is nothing but a visual illusion, because cinema is nothing but a sequence of still pictures shown in rapid succession. Visual illusions in the form of *motion blurring* also find application in the world of modern videogames. Motion blurring is used to reduce the discomfort level of the viewer when the frame-rate of a videogame becomes very low, which can occur because of heavy computational load. At low frame-rates a videogame appears "choppy" or "jerky" which causes discomfort to the player. Motion blurring artificially introduces a blurring effect depending upon the motion in the scene. This enhances the appearance of continuity in the video. The frame-rate of 24 frames per second used in traditional cinema is considered very low by experienced videogame players, who prefer a minimum frame-rate of 60 frames per second. In fact traditional cinema appears to be smooth precisely because it already incorporates the motion blurring effect from its very start in history. Any experienced still photographer will attest that an exposure time of $1/24^{th}$ of a second is too long a duration for producing sharp still images because objects in everyday life move significantly within that time. Still photographers prefer to stay below $1/60^{th}$ of a second. However a motion picture made out of a sequence of very sharp images played back at 24 frames per second will have the aforementioned appearance of choppiness or jerkiness and cause discomfort to the viewer.

Another area of application of visual illusions is in the area of lossy image and video compression, i.e. the efficient representation of visual

data for efficient transmission and storage. Lossy compression techniques reduce data size by encoding the data in such a way that discarding some parts of the data will be imperceptible to the viewer. For example, the higher frequency components of an image can be represented using lower precision numbers than the low frequency components. The introduction of additional noise as a result of this reduction of precision goes unnoticed as it gets lost amid the already high intensity variations at high frequencies. But there do exist frequently occurring situations in which this scheme fails noticeably by introducing conspicuous artifacts into the image. This highlights another practical use of computational models of perception. Models of visual perception can help us in building efficient video and image compression techniques by avoiding such artifacts. Perceptual models of hearing have already been very successfully applied to achieve huge efficiencies in audio data compression. Perceptual models are also useful in image processing where they can be used to reduce perceived noise as opposed to real noise. Other applications exist in watermarking and visual information camouflaging.

1.2 Brightness illusions

Brightness illusions are visual illusions in which different surfaces at different locations of the stimulus presented before a person, having equal physical luminance, are perceived to be of different brightness. This difference in perceived brightness depends on the structure of the stimulus presented. When the areas surrounding the surfaces are covered up, the surfaces appear equally bright again. Note that this difference is not because of the surfaces being projected at different locations of the retina. The nature of the illusion does not change upon sideways shifting of the stimulus. Usually such stimuli are viewed on a display device which have their own light source, but if the illusion is viewed under

reflected light, such as a paper printout, it must be ascertained that all parts are uniformly illuminated in order to ensure that all equiluminant surfaces on screen are also equiluminant on paper, i.e. brightness variations must not be a result of differential illumination. A very famous example of this type of illusion is the Checkershadow Illusion created by Edward H. Adelson, shown in Figure 1.1.

The existence of such illusions implies that the perceived luminance of a region is not determined only by the actual luminance of that particular region. Instead the perceived luminance can be modulated by the observer's brain depending on the area surrounding the region concerned. This illusion is so surprising because most people believe that the perceived luminance of an area depends exclusively on the properties of that area without being affected by the mere presence of any pattern in its surroundings. Henceforth in this work the term intensity shall be used to refer to the actual luminance of a region and the term brightness shall be used to refer to the perceived luminance of that region (though in the field of brightness perception this is also sometimes termed as lightness). In the next section several other brightness illusions are discussed in greater detail.

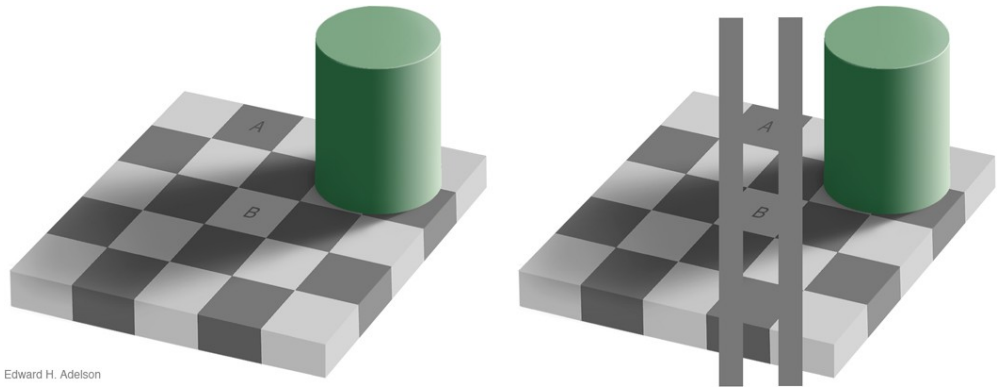


Figure 1.1: The Checkershadow illusion by Edward H. Adelson. In the figure on the left the squares marked A and B are of equal luminance, yet square A clearly appears darker than square B. The figure on the right proves that the two squares are equiluminant by connecting them with solid bars of uniform luminance. (Copyright Information: Edward H. Adelson has allowed free reproduction and distribution of this image as mentioned in http://persci.mit.edu/gallery/checkershadow/download)

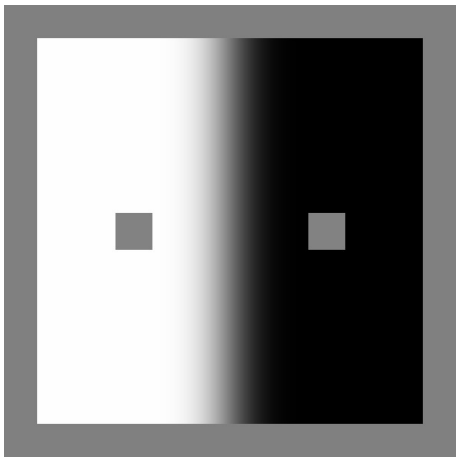
1.3 Examples of brightness illusions

The Checkershadow illusion introduced in the previous section, although strikingly surprising to look at, is too complex to even begun to be analyzed and modeled, due to it being a 2D representation of a 3D scene that consists of lighting-shading effects, shadows, occlusion etc, that increase its complexity. Fortunately brightness illusions can also be seen in far simpler contexts with fewer description parameters which must be modeled and explained before even attempting an explanation of the Checkershadow illusion.

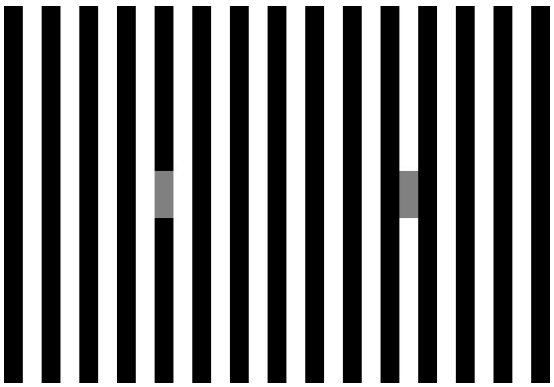
Figure 1.2 and 1.3 show several examples of brightness illusions. The stimuli shown in Figure 1.2 contain two distinct grey patches, usually termed as test patches, similar to the two squares A & B in the Checkershadow illusion of Figure 1.1. These two grey patches have been designed to be equiluminant (by assigning them the same RGB values), yet just like in the Checkershadow illusion the two grey patches appear differently luminant as a result of the surrounding context of those two grey patches. Examples of this include the Simultaneous Brightness Contrast (SBC) (Figure 1.2a), the White's effect (Figure 1.2b), the Shifted-White's effect (Figure 1.2c) and the Checkerboard illusion (Figure 1.2d).

The stimuli in Figure 1.3 show a few other types of brightness illusions which do not provide for any comparative study of two equiluminant test patch perceptions, unlike the stimuli in Figure 1.2.

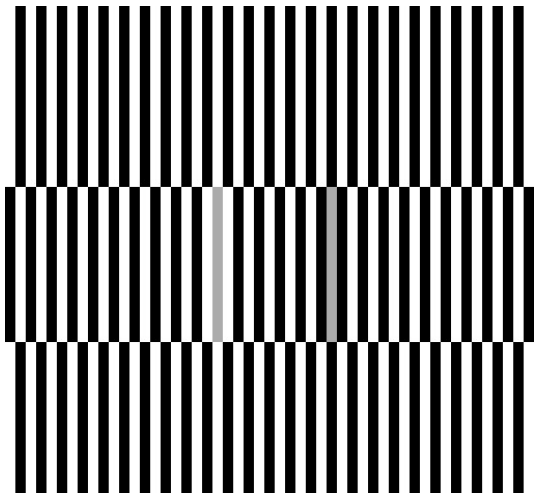
Instead, these stimuli are characterized by a spatially continuous intensity field whose perceived brightness field does not agree with its designed intensity field.



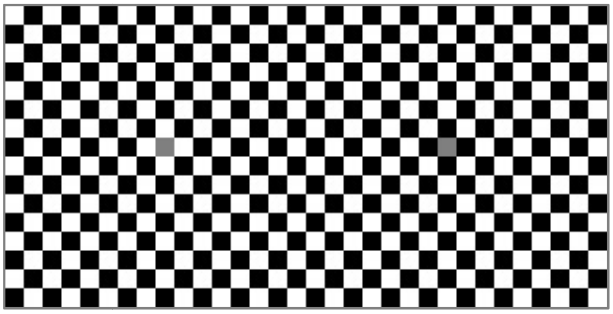
(a) Simultaneous Brightness Contrast.



(b) White Effect.



(c) Shifted White Effect.



(d) Checkerboard illusion

Figure 1.2: Examples of brightness induction illusions (a) Simultaneous brightness contrast (SBC) illusion: The two grey squares have the same real intensity but the square on the dark back- ground apparently looks brighter than the square on the light background. (b) White effect: The two grey bars have the same intensity but the one on the black stripe looks brighter than the one on the white stripe. The bar on the black stripe shares a greater boundary with its two neighbouring white stripes than with the black stripe. Similarly the bar on the white stripe shares a greater boundary with its two neighbouring black stripes than with the white stripe. So, unlike cases a above, here the brightness changes in a direction same as most of its surroundings. (c) Shifted White effect: The two grey bars have the same intensity but the one surrounded on all sides by black regions looks brighter than the one on surrounded by white. So, unlike case a above, here the brightness changes in a direction same as most of its surroundings. (d) Checkerboard illusion: the two grey squares have the same intensity but the one surrounded by white squares looks brighter than the one surrounded by the black squares. The brightness shift is in the same direction as the surrounding brightness.

1.3.1 Simultaneous brightness contrast (SBC)

Figure 1.2a shows a SBC stimulus [Heinemann, 1955]. Two grey squares of equal intensity are drawn over a background such that one grey square is completely surrounded by a uniformly black region and the other grey square is surrounded by uniformly white region. The two grey squares, although having equal intensity, appear unequally bright. From Figure 1.2a, it can be seen that the grey square which is surrounded by white appears darker than the grey square surrounded by black. This means that the direction of brightness change of the grey squares is opposite to that of the respective surrounding regions. In other words the brightness shifts occur in a direction so as to enhance the contrast of the grey squares relative to their surroundings. In the next chapter we shall discuss in detail this phenomenon of brightness-contrast as one of the typical variants of brightness perception complexities.

1.3.2 White's illusion

Figure 1.2b shows White's stimulus [White, 1979]. Two grey patches of equal intensity are drawn over vertically oriented strips of alternating black and white colors. One grey patch is placed over a black strip while the other is placed over a white strip. The two grey patches have equal intensity but appear unequally bright. It can be seen that the brightness of the grey patch sandwiched between two white strips is higher than the brightness of the grey patch sandwiched between the black strips. It should also be noted that the grey patches share a longer border with the strips on their left and right than with the strip they were placed upon and yet the brightness shift of the grey patches occur in a direction so as to enhance the contrast with respect to the strip they share a shorter border with. This shows that the White's illusion is of a completely different nature compared to the SBC illusion. The discovery of White's effect

[White, 1979] was a paradigm shift in the field of Brightness perception because it contradicted the predominant theories of brightness perception that existed up to that time and which were able to provide explanations (qualitative as well as quantitative) for Simultaneous Brightness Contrast.

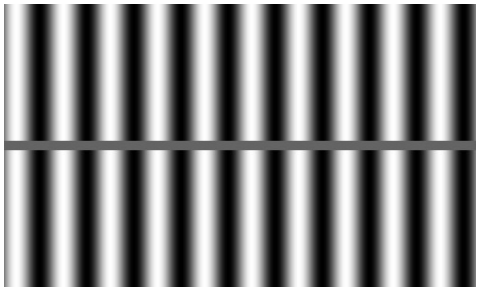
1.3.3 Shifted-White's illusion

Figure 1.2c shows the Shifted-White's stimulus [White, 1981]. This is similar to White's illusion, except that the central region between two imaginary horizontal lines above and below the grey patches is horizontally shifted by a distance equal to the strip width. After shifting, the grey patches are now bordered by either white on all four sides or black on all four sides. The two grey patches have equal intensity but appear unequally bright. It can be seen that the brightness of the grey patch bordered by white regions is higher than the brightness of the grey patch bordered by black regions, i.e. contrast gets reduced as in White's effect. But in the following chapters we shall see that the Shifted-White effect is an even more complex phenomenon compared to the White effect, since depending upon the length scale of the stimulus the brightness shift illusion may even get reversed.

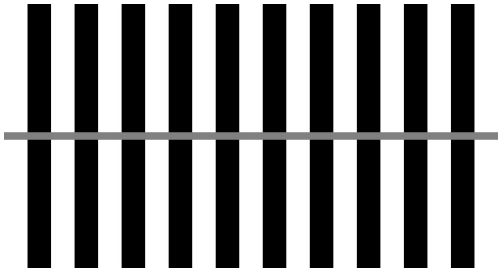
1.3.4 Checkerboard illusion

The checkerboard illusion [De Valois and De Valois, 1988] consists of a chessboard like array of alternating black and white squares as shown in Figure 1.2d. Two squares out of this array, one black and one white, are painted over with the same grey color of uniform intensity. The grey square which was originally black is now bordered on all four sides by white squares and the grey square that was originally white is now bordered by black squares on all 4 sides. It can be observed that the grey square which is surrounded by white squares appears brighter than the

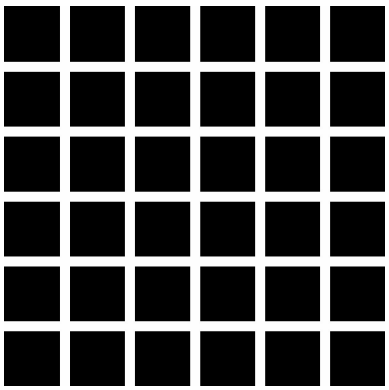
grey square which is surrounded by black squares. The remarkable thing to note here is that the direction of brightness change in this illusion is in a direction so as to reduce the contrast with respect to the bordering squares. This is exactly opposite of what was occurring in the Simultaneous Brightness Contrast (SBC) illusion of Figure 1.2a, where the brightness of the grey squares changed in a direction opposite to the surrounding brightness. The further complexities associated with this illusion will be discussed subsequently in the following chapters.



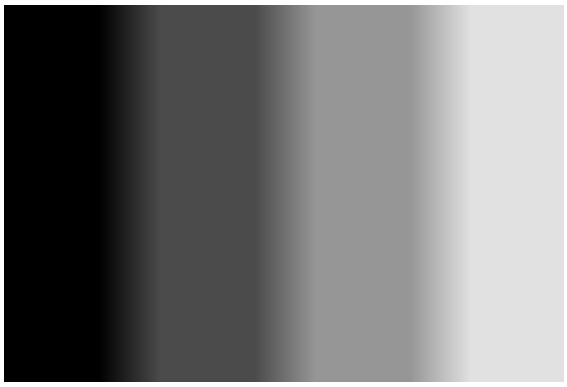
(a) Sine Grating Induction.



(b) Square Grating Induction.



(c) Hermann Grid.



(d) Mach Bands.

Figure 1.3: Further examples of brightness induction illusions (a) Sinusoidal Grating induction: the grey strip has uniform intensity but when placed over a sinusoidally undulating intensity field, appears to have undulating brightness. Here the grey-patch brightness changes in a direction opposite to most of its surroundings considering each of the black or white columns on which it lies. (b) Square Grating induction: the grey strip has uniform intensity but when placed over a square waveform intensity field, appears to have undulating brightness. Similar to the Sinusoidal grating illusion above the grey-patch brightness changes in a direction opposite to most of its surroundings considering each of the black or white columns on which it lies. As described later, we have found through experimental studies that the sinusoidal grating illusion is much stronger in effect than the square grating illusion. (c) Hermann grid illusion: A number of black squares are placed in a grid formation. Splodges of grey are perceivable on the white intersection points, even though the background is purely uniform intensity. (d) Mach Band Illusion: This consists of steps of intensity-plateaus separated by intensity-gradients. Bright and dark bands can be observed along the lines where the gradients meet the plateaus. These bright and dark bands are illusory brightness-peaks and brightness-troughs, respectively, in the brightness profile, as the input stimulus has a monotonic intensity profile devoid of any peaks or troughs.

1.3.5 Grating induction illusion

The grating induction illusion [McCourt, 1982] (Figure 1.3a and 1.3b) consists of a grating of alternating black and white wave-like pattern, upon which a grey colored strip of uniform intensity is drawn over. This

grey strip appears to have undulating brightness so as to enhance the contrast with respect to its local neighbourhood. There exist two separate versions of this illusion as described below. Although the two types may seem to be only slight variants, the illusory effect of one is much stronger than the other.

1.3.6 Sinusoidal grating induction

Figure 1.3a shows a sinusoidal grating induction stimulus. A single grey strip of uniform intensity runs horizontally across a sinusoidally varying background intensity field. The grey strip appears to have an undulating brightness. Careful examination will reveal that the grey strip appears bright at background intensity minima and it appears dark at background intensity maxima. Hence the perceived brightness change is enhancing contrast.

1.3.7 Square grating induction

Figure 1.3b shows a square grating induction stimulus. A single grey strip of uniform intensity runs horizontally across a background intensity field varying as a square waveform. The grey strip appears to have an undulating brightness. Careful examination will reveal that the grey strip appears bright where the background is dark and appears dark where the background is bright. Hence the perceived brightness change is enhancing contrast. This illusion might superficially seem to be same as the sinusoidal grating induction illusion, but we have been able to show through our experiments in the subsequent chapters that the sinusoidal grating induction is a much stronger effect than the square grating induction.

1.3.8 Hermann grid illusion

Figure 1.3c shows the Hermann grid stimulus, consisting of a uniformly spaced grid of black squares separated by white lanes as in the case of evenly spaced city blocks. Grey blobs are visible at the lane intersections although the lanes themselves are of uniform white intensity. Compared to all the previous illusions the Hermann grid illusion has one remarkable difference, which is that when one looks carefully at one lane intersection one can easily verify that it has uniform intensity, but when one ceases to concentrate on a single intersection, the grey blobs reappear. In contrast, in all of the previous illusions the illusory effect remains the same no matter whether the observer is concentrating or not. In fact an observer will not even notice that the intensities of the grey patches were the same and that any apparent difference was only illusory, if he is not specifically told so by another person. In the Hermann grid illusion any alert observer would immediately notice that the grey blobs are illusory. In a later chapter we shall discuss several variations of the Hermann grid illusion and some of its properties.

1.3.9 Mach band illusion

Figure 1.3d shows an example of the Mach band illusion. The Mach band illusion occurs at the boundary between two regions of differing intensity gradients. The border along which these two regions meet shows either a brightness-peak or a brightness-trough. These peaks (bright Mach bands) and troughs (dark Mach bands) are of illusory nature. A bright Mach band can be seen where the gradient is decreasing and a dark Mach band can be seen where the gradient is decreasing. Figure 1.4b shows the horizontal intensity profile of the Mach band stimulus shown in Figure 1.4a. The term intensity profile refers to a graphical plot of intensity versus distance of an image when one moves along a straight line through the image. The horizontal intensity profile is the intensity profile along a

horizontal line through the image. Similarly the vertical intensity profile is the intensity profile along a vertical line through the image. In the case of both Figure 1.3d and Figure 1.4a the horizontal intensity profile is the same along all horizontal lines through the image. This may not be true in other contexts, as for example Figure 1.2, and in such cases it shall be specified which horizontal line is used to plot the horizontal intensity profile. In most cases (but not always) it is a horizontal line passing through the center of the image.

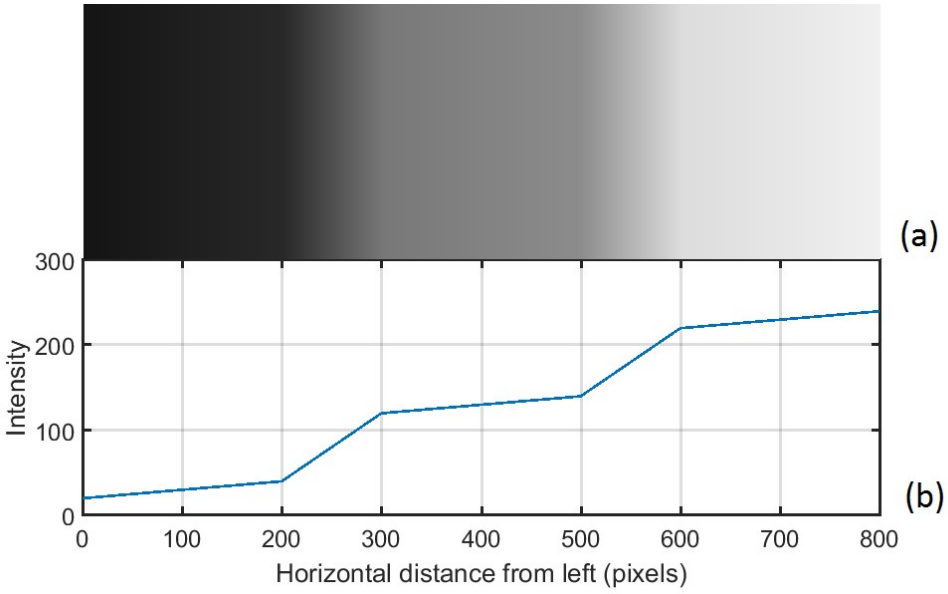


Figure 1.4: (a) Mach band stimulus image. (b) Horizontal brightness profile for the Mach Band stimulus.

Figure 1.4b shows the intensity profile of a Mach band stimulus juxtaposed alongside the stimulus itself. It is clear that the intensity profile is a non-decreasing function of position. Yet when looking at Figure 1.4a one can clearly see brightness-peaks at the positions where the gradient decreases abruptly. One can also see brightness-troughs at the positions where the gradient increases abruptly. But there can exist no peak or trough in a non-decreasing function. The bright peaks and dark troughs

must therefore be of illusory origin. These peaks and troughs appear all along the borderline wherever the gradient changes abruptly, giving them the look of elongated bands. Hence the name Mach bands, in the honour of Ernst Mach who first observed them. We shall explore several other properties of the Mach band illusion in some of the subsequent chapters.

In the following chapter we shall review various models and scientific theories of brightness perception based on experimental research involving many of the illusions discussed in this chapter.

Chapter 2

History, existing models and theories

2.1 Introduction

The first formal studies of perceptual phenomena were made by Gustav Theodor Fechner (1801-1887), who is now known as the founder of the field of Psychophysics. He demonstrated that perception could be systematically measured and modeled.

In his publication Elemente der Psychophysik [Fechner, 1860], Gustav Theodor Fechner first introduced a class of techniques and psychophysical methods in order to correlate stimulus and sensation [Gescheider, 1997] [Ehrenstein and Ehrenstein, 1999]. Fechner realized that psychophysics requires the objectification of subjective response. This task can involve 'detection', 'identification' or 'discrimination' of various stimuli. He proposed methods for psychophysical experimentations which involve measurement, quantification and characterization of the perceptual experience using some psychophysical quantities. One of the important measurable quantities to judge any stimulus and objectify its perception in brain is the Absolute threshold of a stimulus sensation. The term 'Absolute threshold' is defined as the minimum intensity level of stimulus that can be detected reliably by any subject. Fechner continued the earlier work of E.H. Weber, one of whose important findings were, if a stimulus of intensity I is applied on one of the senses, then the minimum amount ΔI that the applied intensity must be increased in order for the change to be perceptible, is proportional to the applied intensity I. The ratio $\Delta I/I$ is called the Weber fraction. Fechner added on to Weber's work which led to the Weber-Fechner Law which states that, "In order that the intensity of a

sensation may increase in arithmetical progression, the stimulus must increase in geometrical progression". This implies that the perceived strength of a given sensation is a logarithmic function of the intensity of the stimulus which generated that sensation. This probably is the earliest example of an attempt to mathematically model perceptual experience. In the field of visual perceptual experience, the Mach Band illusion demonstrated in the previous chapter can be said to be the first brightness illusion to be scientifically studied. It was the Austrian Physicist [Ratliff, 1965] Ernst Mach (1838-1916) in 1865 who used rapidly rotating drums with varying amounts of black and white regions painted over it for producing any shade of grey in reproducible and measurable amounts with any desired intensity gradients. As the drums rotate rapidly regions with more white than black would have more intensity in direct proportion to the amount of white. This led Mach to identify those bright and dark bands, now known as Mach bands, as a means to learn about the internal mechanisms of the visual processing system. Mach bands are not only present in laboratory or artificial situations, they may easily be observed at the edge of practically all shadows where light or dark lines will surround the penumbra. Fomm's striae [Fomm, 1896], seen while determining the wavelength of X-ray from diffraction experiments, turned out to be nothing but results of Mach band illusions and a serious mistake in experimental physics [Wind, 1899] was thus identified and corrected. This brightness perception illusion was also found to be the culprit in the well-known discrepancy in determination of Earth's radius from its shadow during lunar eclipse and the correct explanation was finally provided by physiological/perceptual optics rather than by physical or geometrical optics. Despite having such a long historical presence in the field of visual perception, even to this day Mach bands still remain an excellent subject of study in linking perception with the underlying neural

mechanisms and are continuing to raise intriguing research problems as we shall see in Chapter 7 of this thesis.

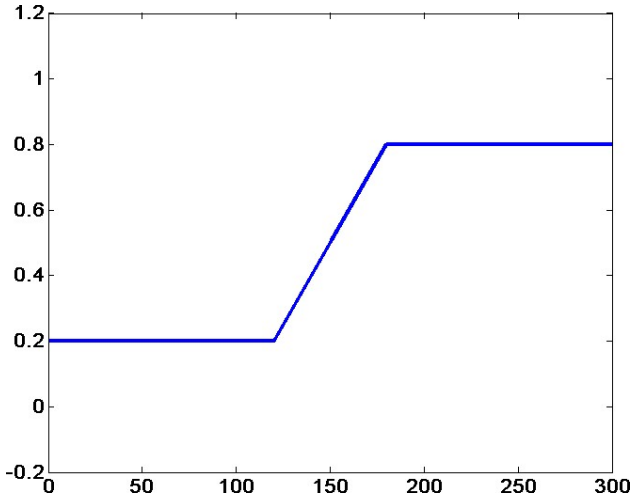
Mach himself proposed a spatial filtering model for explaining Mach bands using the Laplacian operator. When the Laplacian operator is applied to an edge, it produces a very strong positive response on the higher side of the edge and a very strong negative response on the lower side. It also produces zero response for uniform regions and for uniform gradients. When the Laplacian image is subtracted from the original stimulus it produces an effect similar to the Mach bands, generating a brightness-peak where the gradient abruptly decreases and a brightness-trough where the gradient abruptly increases. If the functions *u(x, y)* and *v(x, y)* denote the input stimulus and the perceived brightness fields respectively, then Mach's model can be written as,

$$v(x,y) = u(x,y) - m * \nabla^2 u(x,y) \qquad (2.1)$$

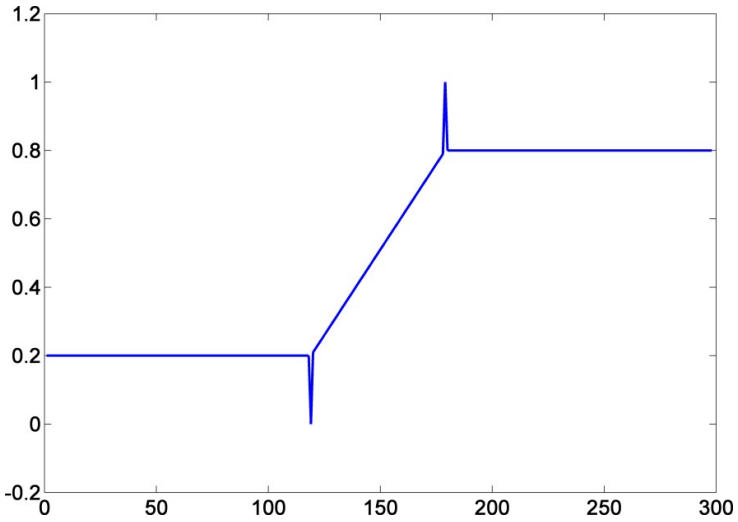
where m is a positive constant which determines the strength of the Mach bands.

The response curve of this equation for a typical Mach band image is shown in Figure 2.1b.

Mach's model turned out to have various shortcomings, which shall be discussed later, but it was the first attempt of what are called spatial filtering models of brightness perception.



(a) Input profile



(b) Output profile

Figure 2.1: (a) Input brightness profile for a typical Mach Band stimulus. (b) Output profile produced by Mach's model in Equation 1.1. The thin peak and trough on the right and left respectively represent the light and dark Mach bands.

2.2 Spatial filtering models

A spatial filtering model takes the 2D intensity field of the stimulus as input and performs a single or a series of spatial filtering operations on it,

such as 2D convolution with a kernel function, and produces an output field, which is interpreted as perceived brightness values. The spatial filter need not necessarily be a linear filter. For example, after performing a sequence of linear operations, the individual outputs may be combined later using a non-linear operation, which would make the whole operation non-linear. This includes the so far most successful spatial filtering model in brightness modeling, namely the ODOG filter (ODOG stands for Oriented Difference of Gaussians) of Blakeslee and McCourt [Blakeslee and McCourt, 1999] [Blakeslee and McCourt, 2004] [Blakeslee et al., 2016] [McCourt et al., 2016].

2.2.1 Lateral inhibition

Spatial filtering models have been supported by the experimental observation of lateral inhibition in the retina of the eye as well as the primary visual cortex [Rodieck and Stone, 1965] [Hubel and Wiesel, 1959]. Lateral inhibition is the phenomenon in which the response of a neuron receiving a stimulus is inhibited by the activity of a neighbouring neuron which is also receiving a stimulus. Lateral inhibition has been experimentally observed in the retina and the LGN of organisms. Lateral inhibition makes neurons more sensitive to spatially varying stimuli than to spatially uniform stimuli. This is because a neuron getting stimulated by a spatially uniform stimulus is also inhibited by its surrounding neurons, thus suppressing its response. On the other hand a neuron subjected to a spatially varying stimulus is less inhibited by its neighbours that are not excited, thus producing stronger response. Therefore in the case of visual neurons, lateral inhibition makes them more sensitive to edges in the scene. Although usually described for visual neurons, lateral inhibition is also found in other sensory systems, such as auditory and olfactory neurons. The region on the retina to which a ganglion cell

responds is called its receptive field. A typical receptive field, termed as an on-centered receptive field, consists of a central *excitatory* region surrounded by an *inhibitory* region. When light falls on the excitatory region it causes an increase in the output response of the ganglion cell, hence the term excitatory. On the other hand when light falls on the inhibitory region it decreases the output response of the ganglion cell, hence the term inhibitory. There also exist off-centered receptive fields in which the inhibitory region lies at the center which the excitatory region surrounds. Next we shall see how to mathematically model such excitatory/inhibitory behaviour.

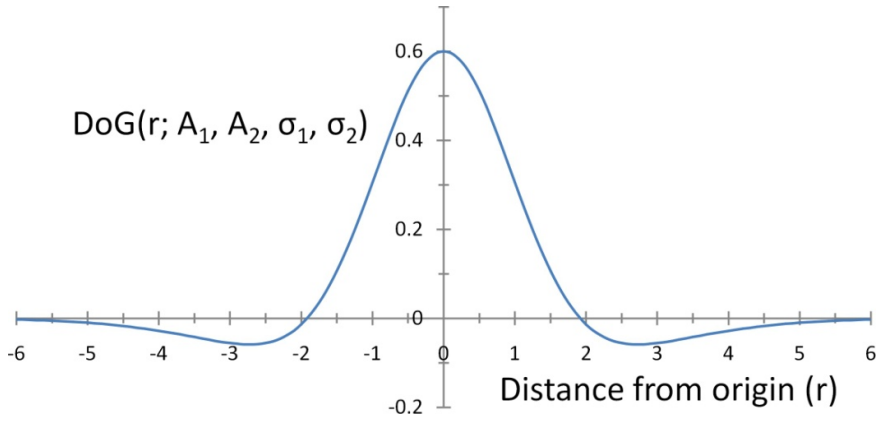


Figure 2.2: Typical shape of the DOG function. The horizontal axis represents the spatial distance r from the center of the receptive field. The y-axis is the response produced at the origin by a spot of light falling at a distance r from the origin.

2.2.2 DOG model

A very well known spatial filtering function that has been frequently applied to model this phenomenon of lateral inhibition is the DOG (Difference of Gaussian) function defined as:

$$DOG(r; A_1, A_2, \sigma_1, \sigma_2) = A_1 e^{\left(\frac{r^2}{2\sigma_1^2}\right)} - A_2 e^{\left(\frac{r^2}{2\sigma_2^2}\right)} \quad (2.2)$$

where, r is the radial distance from the center of the receptive field. The parameters σ_1 and σ_2 signify the widths of the two Gaussian functions. In order to mimic lateral inhibition we must have an on-centered receptive field, i.e. on whose excitatory region is at the center and an inhibitory region in the surround, i.e. the filtering function must be positive for small values of r and negative for larger values of r, and finally converging to zero for very large values of r. This can be achieved when we choose $\sigma_2 > \sigma_1$ and $A_1 > A_2 > 0$. Under such conditions the DOG function cross-section looks as shown in Figure 2.2. The first Gaussian, having a narrower width contributes more to the central positive region whereas the second wider Gaussian contributes more to the surrounding negative region. The DOG function therefore qualitatively mimics a typical receptive field.

Apart from an on-centered receptive field one can also have an off-centered receptive field, i.e. one with inhibitory center and excitatory surround. Both on-centered and off-centered neurons are known to be present in the retina and the LGN, which is a structure in the thalamus of the brain.

Although the DOG filter can account for the Simultaneous Brightness Contrast illusion (SBC), it cannot explain the White's effect along with the Shifted-White and Checkerboard illusions, which were demonstrated in the previous chapter. The visual perception community is still in search of a single linear filtering algorithm that can simultaneously account for the various types of brightness illusions demonstrated in Figures 1.2 & 1.3 of Chapter 1.

2.2.3 The challenge in modeling brightness induction

As mentioned in section 1.3.2 the nature of White's effect is very different to that of Simultaneous Brightness Contrast. A careful look at Figure 1.2

will reveal that they can be broadly categorized into two contrary types of brightness illusions which may be termed as the *brightness-contrast* and *brightness-assimilation* types of illusions. In the brightness-contrast type, the apparent brightness of a region shifts in a direction so as to enhance the contrast with respect to its surrounding regions. In other words the perceived luminance gets modulated opposite to the direction of the surrounding regions. An example of this is the Simultaneous Brightness Contrast (SBC) illusion (Figure 1.2a). It may also be noted from the figures 1.3a & 1.3b that the Grating induction illusions are also of brightness-contrast type. In the brightness-assimilation type of illusions, the apparent brightness shifts in the same direction as its surroundings, as if it were assimilating the intensity of its surroundings. Examples of this include the White's effect (Figure 1.2b), the Shifted-White's effect (Figure 1.2c) and the Checkerboard illusion (Figure 1.2d). This process of a given surface's brightness either shifting towards or shifting away from its neighbouring surfaces' brightness is called *brightness-induction*, as if the brightness shift of the grey patch is being induced by the surrounding regions. This is why the visual perception community is looking for a universal model that can simultaneously explain this bi-directional induction of brightness.

2.2.4 ODOG, LODOG, FLODOG models

The Oriented Difference of Gaussians (ODOG) model by Blakeslee and McCourt [Blakeslee and McCourt, 1999] [Blakeslee and McCourt, 2004] [Blakeslee et al., 2016] [McCourt et al., 2016] was proposed as an extension of their previous DOG model [Blakeslee and McCourt, 1997]. This spatial filtering model was capable of explaining a large number of illusions. This model was very successful in being able to explain both brightness-contrast and brightness-assimilation illusions using the same spatial filtering algorithm. Usage of anisotropic filters and nonlinear

normalization are two key features which make the ODOG model different in nature from its predecessors. This model utilizes 42 different ODOG filters in total. A single ODOG filter can be expressed as:

$$ODOG(x, y; \sigma_f) = \frac{1}{2\pi\sigma_f^2} e^{\frac{-x^2}{2\sigma_f^2}} \left[e^{\frac{-y^2}{2\sigma_f^2}} - \frac{1}{2} e^{\frac{-y^2}{8\sigma_f^2}} \right] \quad (2.3)$$

Along the major axis of any single ODOG filter, the surround Gaussian of the filter has twice as much spread as the central Gaussian. But in the orthogonal direction to the major axis (minor axis), both the centre and surround extend by same magnitude (Figure 2.3).

Figure 2.3: Typical shape of the ODOG function at a particular scale and orientation.

Seven ODOG filters, whose standard deviations σ_f form a geometric series ranging from $1.5/\sqrt{2}$ to $96/\sqrt{2}$ with common ratio 2, as defined by equation 2.4, are then added to form the Multiscale ODOG filter function as defined in equation 1.5:

$$\sigma_f = \frac{3}{2\sqrt{2}} 2^f; f = \{0,1,2,3,4,5,6\} \quad (2.4)$$

$$Multiscale\_ODOG(x, y) = \sum_{f=0}^{6} A_f \times ODOG(x, y; \sigma_f) \quad (2.5)$$

where,

$$A_f = C \times {\sigma_f}^{-0.1}; f = \{0,1,2,3,4,5,6\} \quad (2.6)$$

Next we rotate the *Multiscale_ODOG* filter function as in Equation 2.5 above through an angle $\boldsymbol{\theta_i}$, using equation 2.7, where $\boldsymbol{\theta_i}$ are assigned six equally spaced values in the interval [0, π]. For all six $\boldsymbol{\theta_i}$ we obtain six *Rotated_Multiscale_ODOG* filter functions; one filter per orientation (see Figure 2.3).

$$Rotated\_Multiscale\_ODOG(x, y, \theta_i) = Multiscale\_ODOG\left(\begin{bmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\right) \quad (2.7)$$

with

$$\theta_i = \frac{\pi}{6} \times i; i = \{0,1,2,3,4,5\} \quad (2.8)$$

Each of the 6 rotated multiscale ODOG functions are then individually convolved with the input stimulus image *I(x,y)* to produce 6 intermediate oriented responses $R(x, y; \theta_i)$, defined by:

$$R(x, y, \theta_i) = I(x, y) * Rotated\_Multiscale\_ODOG(x, y, \theta_i) \quad (2.9)$$

where $*$ represents the 2D convolution operation.

The final output is then generated by a non-linear combination of these 6 inter mediate directional responses by normalizing each oriented response by its RMS value:

$$FINAL\_ODOG\_RESPONSE = \sum_{i=1}^{6} \frac{R(x,y,\theta_i)}{RMS[R(x,y,\theta_i)]} \quad (2.10)$$

In the ODOG model implementation by Robinson et al. [Robinson et al., 2007a], which we have used in this thesis for comparing various spatial filtering models including the models proposed in this thesis, the values of the centre frequencies (in cycles/degree) of the ODOG filters and their corresponding weights, as used by Robinson, are tabulated in Table 2.1.

Table 2.1: Weights used for each of the 7 center frequency values in Alan Robinson's [Robinson et al., 2007a] implementation of ODOG

Center frequency (σ_f)	0.1	0.2	0.4	0.8	1.6	3.2	6.4
Weight (A_f)	0.79	0.85	0.91	0.98	1.05	1.12	1.20

Alan Robinson has further extended the ODOG model in two ways, referred to as LODOG and FLODOG [Robinson et al., 2007b]. LODOG, which stands for *Locally normalized ODOG*, modifies the final RMS normalization step (equation 2.10) of ODOG by performing a localized RMS normalization at each point using a small window of fixed size. FLODOG, which stands for *Frequency-specific Locally normalized ODOG*, extends this even further by doing a localized RMS normalization at each point for each frequency and uses a window size which is appropriate for that frequency. The high frequency, i.e. small σ_f ODOG responses use small windows, whereas low frequency ODOG responses use large windows for their RMS normalization steps.

In this thesis we shall mostly not concern ourselves with LODOG and FLODOG. While LODOG is a bit more computationally intensive than ODOG, the FLODOG is much more so. A major part of this thesis is about explaining brightness illusions using models of low complexity. Besides, there exist certain other limitations of this class of spatial filtering models which we shall elaborate in Chapter 3 of this thesis. Apart from spatial filtering models there exist several other approaches to modeling of brightness perception. The important ones among them are the *Gestalt-anchoring* models and *Edge integration* models which are described next.

2.2.5 Gestalt-anchoring models

The Gestalt school of thought in psychology proposes that perception can only be understood holistically and not as a sum of its parts, i.e.

perception cannot be explained by analyzing it into small parts, studying the behavior of each part and then reconstructing the behavior of the whole by joining together these parts. Although this reductionist approach has been very successful in other branches of science, such as physics and chemistry, Gestalt psychologists do not believe this approach would work for understanding perception or even the workings of the brain in general.

The term *anchoring* refers to the approach in which certain intensity levels in the input stimulus are assigned (anchored) to particular perceived brightness levels and the other brightness levels are determined by these anchored levels.

In gestalt anchoring models the anchors are determined by gestalt grouping principles, i.e. the anchor levels are determined from the stimulus as a whole and not determined locally for every point and its local surroundings. In other words the brightness levels are determined not by the local context alone but by the global context (which includes the local context). The most well known anchoring model is that of Gilchrist [Gilchrist et al., 1999] [Gilchrist, 2006]. According to this model, at every context level the highest intensity value is anchored to the perceived brightness level of 'white' (i.e. maximum perceivable brightness) and the rest of the intensity levels are assigned brightness values in proportion to the highest intensity level (which was anchored to white). But this assigned brightness level is only for one context level and is not the final perceived brightness. The final perceived brightness is computed by taking the mean of the perceived brightness at every context level. Note that this theory does not specify what those context levels are and how to find them.

Figure 2.4: SBC stimulus for demonstrating failure of Gilchrist's Gestalt anchoring theory which predicts that the two square grey patches should have equal perceived brightness. We can clearly see that the two grey patches appear differently bright. Both of the grey patches are equiluminant but possess higher intensity than either of their surroundings.

In a complex real world image it would not be clear what the context levels should be. Thus appears the gestalt nature of the theory. The context levels are assumed to have been recognized by some other processes which have to consider the image as a whole. Gilchrist's theory can nicely explain the Simultaneous Brightness Contrast (SBC) stimulus of Figure 1.2a, but fails in the case shown in Figure 2.4 where both the equiluminant patches have higher luminance values than their respective surroundings.

2.2.6 Edge integration models

Edge integration models [Land and McCann, 1971] combine intensity contrasts across the local edges present in an input stimulus, with anchored brightness values (as described in previous Gestalt-anchoring models section). This implies that sharp luminance jumps across edges play a bigger role in determining the final perceived brightness than gradual changes of luminance. The final perceived brightness is computed by spatially integrating the local edge contrasts from all directions. [Grossberg and Todorovic, 1988] [Grossberg et al., 1997].

2.2.7 Thesis contribution

In this thesis we shall investigate brightness perception from the perspective of spatial filtering. In the beginning we shall describe new psychophysical experiments using modified variants of the White and Shifted-White stimuli which demonstrate some weaknesses of the ODOG model, which is till date the pre-eminent model of brightness perception (Chapter 3). Then we shall introduce a new model of brightness perception which is inspired from the visual pathways of the human visual system, apart from being much simpler than the ODOG model (Chapter 4, 5). Next we shall present an application of brightness perception models in noise reduction (Chapter 6). We shall argue that the use of perceptual models increases the effectiveness of noise reduction via image smoothing while preserving edges at the same time. In the last part (Chapter 7) we shall explore the spatial scaling behaviour of brightness illusions using novel variations of Mach Band, Sine Grating, Square Grating and Herman Grid illusions. The experiments in this chapter demonstrate the various complexities of brightness perception modeling mostly untouched in the earlier spatial filtering approaches, as well as in the previous chapters of the present thesis and in fact go on to expose some limitations of spatial filtering models of brightness perception. The contributions of the present thesis are schematically represented in the diagram below.

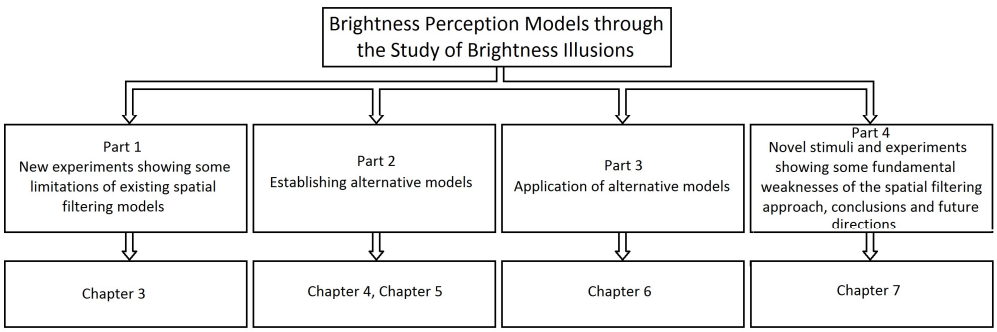


Figure 2.5: Schematic of thesis contribution

Chapter 3

Modeling brightness induction: challenges facing existing models

3.1 Introduction

As late as the year 1999 there was no single spatial filtering model which could simultaneously explain both brightness-contrast and brightness-assimilation (see section 2.2.3), so much so that after the introduction of White's effect in 1980 some researchers started believing that this goal was impossible to achieve by any spatial filtering approach. In 1999 Blakeslee and McCourt proposed the Oriented Difference of Gaussian (ODOG) [Blakeslee and McCourt, 1999] spatial filtering model (see section 2.2.4) which could simultaneously predict the brightness shifts in the Simultaneous Brightness Contrast (SBC), White's effect and Grating Induction illusions. It could also explain some cases of the Todorovic illusion [Todorovic, 1997]. In 2004 Blakeslee and McCourt [Blakeslee and McCourt, 2004] used ODOG to further explain the Shifted-White and Checkerboard illusions. The success of the ODOG model showed that brightness induction could be explained without resorting to any high level interpretation of the scene such as detecting features like T-junctions or detecting objects within the scene [Blakeslee et al., 2016] [McCourt et al., 2016]. ODOG manages to account for both brightness-contrast and brightness-assimilation by only doing low level pixel operations on the image using a combination of several anisotropic spatial filters. The existence of directional edge detection neurons within the visual cortex was already well known because of which such oriented spatial filters are

also biologically justified. However what is not biologically justified is the final global contrast normalization step, as noted by Kingdom [Kingdom, 2011], which requires a global integration of visual signal from all across the visual field (although the LODOG [Robinson et al., 2007a] and FLODOG models [Robinson et al., 2007b] avoid this difficulty by normalizing locally). Notwithstanding these objections Blakeslee and McCourt demonstrated the successful application of their model over a wide range of spatial frequencies and also a wide range of patch sizes.

3.2 Limitations of the ODOG model: new experiments

Even after the aforementioned successes of the ODOG model we have been able to find some of its shortcomings [Bakshi et al., 2016] by designing some new visual stimuli for which the ODOG model predicts the direction of brightness shift which is contrary to experimental findings. These experiments concern variants of the White's and Shifted-White stimuli.

The White and Shifted-White illusions [White, 1979], although both being brightness-assimilation illusions, show some complexities with respect to change of scale. They behave the same way at small scales but switch behaviour at larger scales. We have also found some limitations of the ODOG filter by comparing the ODOG output with the experimental observations of the White and Shifted-White stimuli at various length scales. We compare the ODOG output with experimental observations of the same. The ODOG output shows a more intuitive behaviour, which unfortunately does not match with reality, as it does not switch behaviour with length scales. The experimental details are described below.

White and Shifted-White stimuli at various length scales were generated by varying two parameters, namely, test patch length and

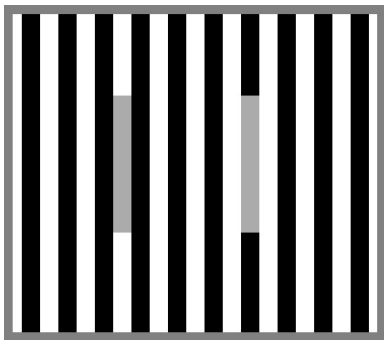
spatial frequency. For the purpose of experimental validation a given stimulus was presented to two test subjects who were told to match the perceived gray-patch brightness from a separate palette of gray scale intensities varying from black to white, that was displayed alongside the stimulus on the same display screen. The final brightness value is calculated by taking the mean of 5 readings obtained from each of the test subjects. The stimuli were displayed as 30.5 degree by 30.5 degree images (viewed from a distance of 48.3 cm) on a high resolution LCD monitor. The monitor had been linearized with the help of a photometer. Subjects were given a rest time between the consecutive illusion readings in order to avoid any after-image of the last stimuli displayed in the screen and readings were taken after having a constant gaze at the stimulus.

3.2.1. White's illusion

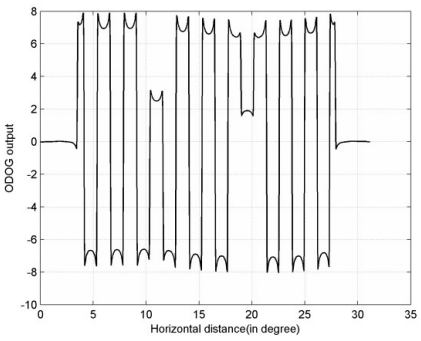
Figure 3.1a shows a White's illusion with a large patch-height of 9.35 degrees (scaled down for illustration). Its ODOG output profile, shown in Figure 3.1b, however predicts that the left grey patch, i.e. the one which is in between the two black strips, should be brighter than its counterpart. But in reality the grey patch on the left, as can be readily seen in Figure 3.1a, appears much darker than the right one and the effect is very prominent and unambiguous. Figure 3.1d shows a plot of the experimental observations of the apparent brightness of the two grey-patches as seen by the subjects, with respect to the patch-height. It clearly shows that the grey-patch in between two white bars (i.e. the grey-patch on the right in Figure 3.1a) always looks brighter than the other grey-patch at all patch-lengths. This is contradicted by Figure 3.1c which is a plot of the ODOG predicted brightness of both the grey-patches with respect to patch-height. In Figure 3.1c the two curves intersect, i.e. there is a specific patch-height beyond which the ODOG predicted brightness gets

inverted. This point is called the Threshold Inversion Point (TIP) from here on.

The TIP is present in the ODOG output of the White effect for a very wide range of spatial frequencies. This is shown in Figure 3.1e, for a patch-height value of 9.35 degrees, i.e. beyond the TIP. In Figure 3.1e the brightness curve for the grey-patch on co-axial black bar is below the brightness curve for the grey-patch on co-axial white bar. Figure 3.1f shows the corresponding experimental curves. It shows that the brightness curve for the co-axial black bar should be above the brightness curve for the grey-patch on co-axial white bar, unlike the results in Figure 3.1e. For comparison the LODOG and FLODOG output profiles for the input stimulus in Figure 3.1a are shown in Figure 3.1g and Figure 3.1h respectively. LODOG, just like ODOG, incorrectly predicts the direction of brightness shift, whereas FLODOG can correctly predict the direction of brightness shift. However, with FLODOG being several times more computationally intensive than either ODOG or LODOG, the problem of explaining brightness illusions with low complexity models still remains.



(a)



(b)

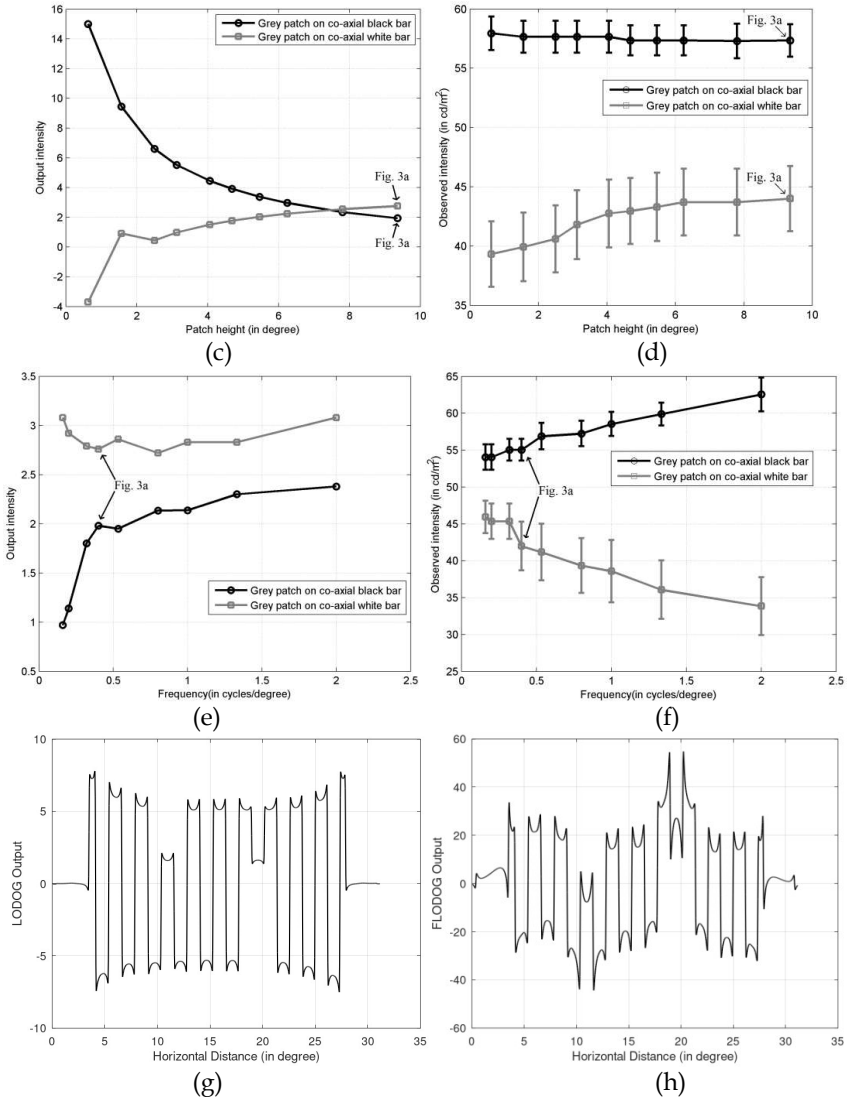


Figure 3.1: (a) Input stimulus of White's illusion with patch-height of 7.9cm (9.35 degrees at 48.3 cm) and frequency 0.47 cycles/cm (0.40 cycles/degree at 48.3 cm) (scaled down to fit in available space) (b) Output intensity graph from ODOG for the stimulus shown in Figure 3.1a incorrectly predicts that the left grey patch should be brighter than the right grey patch. (c) Intensity vs. Patch-height characteristic, as predicted by ODOG at a constant frequency of 0.40 cycles/degree. It can be seen that there is an inversion in intensity characteristic as the patch-height is increased. The point of inversion in intensity is called Threshold Inversion Point (TIP) (7.48 degrees, here). (d) Experimental subject observation of intensity vs. patch-height at a constant frequency of 0.40 cycles /degree. (e) ODOG predicted curves of intensity vs. frequency at a fixed patch-height of 9.35 degrees. (f) Experimentally observed frequency characteristic at a fixed patch-height of 9.35 degrees. (g) LODOG Output profile for the stimulus shown in Figure 3.1a incorrectly

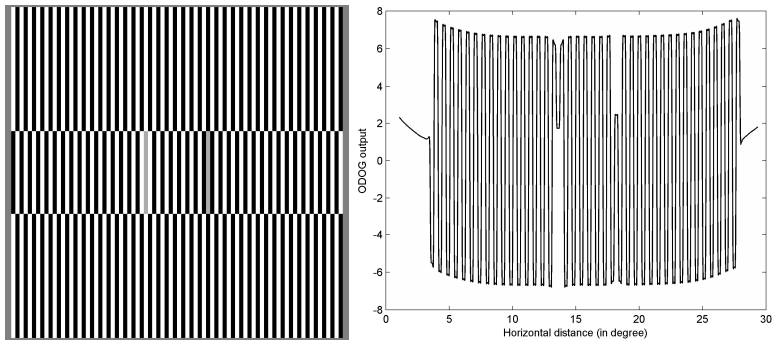
predicts that the left grey patch should be brighter than the right grey patch. (h) FLODOG Output profile for the stimulus shown in Figure 3.1a is able to correctly predict that the left grey patch should be darker than the right grey patch.

3.2.2. The Shifted White effect

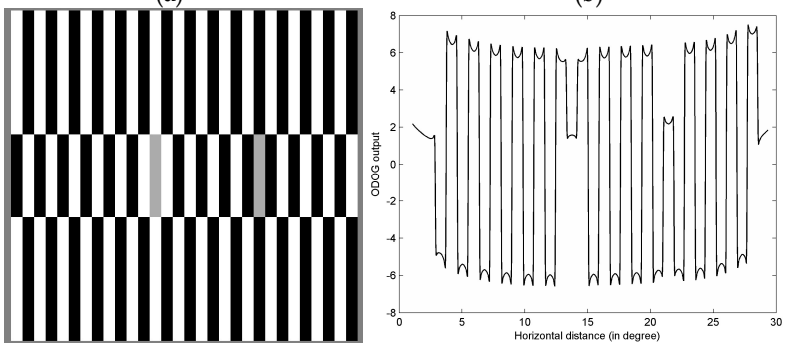
The Shifted White illusion, as shown in Figure 3.2a, is a modified White's illusion where the portion of the grating containing the grey patches are shifted horizontally by a distance of one bar width with respect to their upper and lower regions.

The graph in Figure 3.2g shows the ODOG predicted brightness variation of the two grey patches with respect to patch height. Similar to the case of White effect above, ODOG incorrectly predicts an inversion in the perceived brightness of the two grey patches with increasing patch height, whereas the experimental brightness curves shown in Figure 3.2h show no such inversion. The inversion occurs as a patch height of 4.8 degrees, which we call the Threshold Inversion Point (TIP).

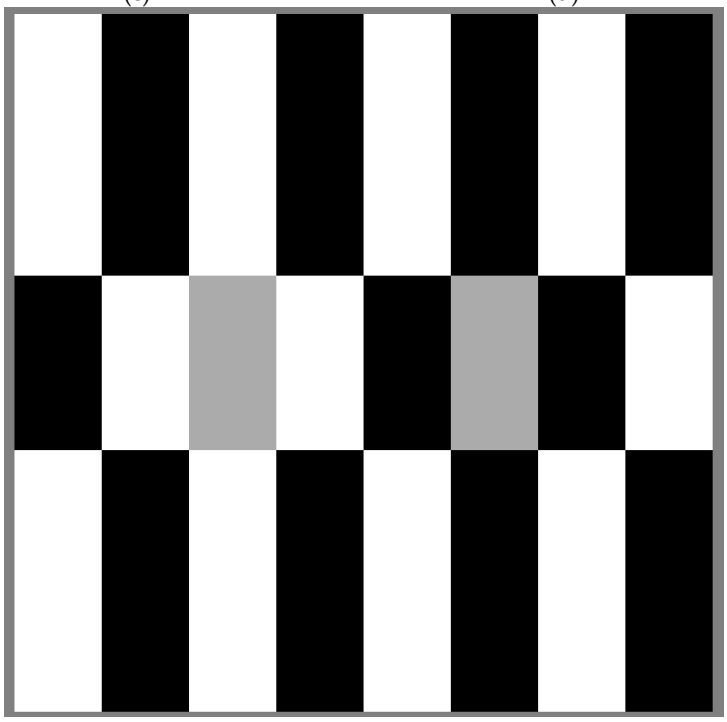
Figure 3.2i shows the ODOG predicted brightness curve of the two grey patches with respect to horizontal spatial frequency. While the corresponding experimentally observed graph in Figure 3.2j shows an inversion point at 0.57 cycles/degree, the ODOG predicted curves show no such inversion. Thus ODOG fails to predict the experimental curves both with respect to patch height and with respect to horizontal spatial frequency.


ODOG output
Horizontal distance (in degree)

(a) (b)


ODOG output
Horizontal distance (in degree)

(c) (d)



(e)

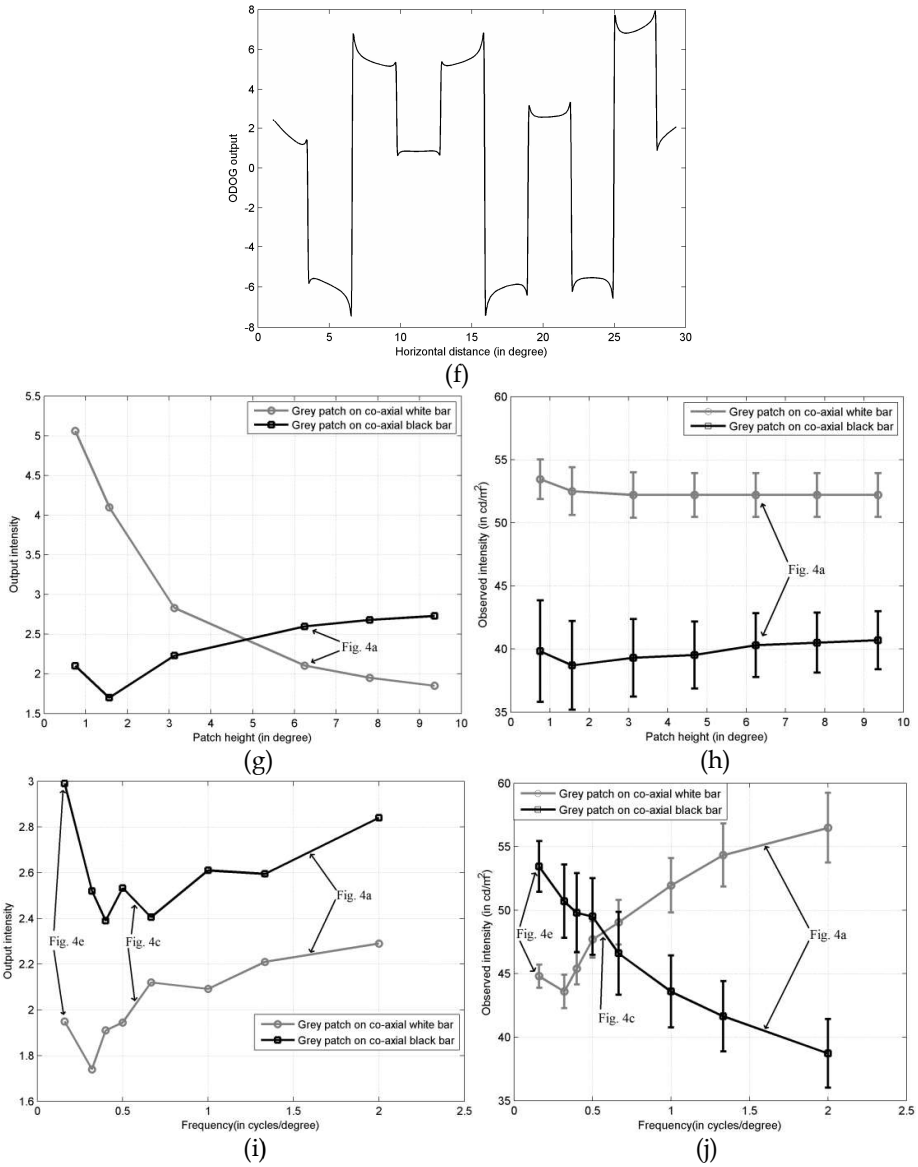


Figure 3.2: (a) Shifted White input stimulus having a patch-height of 5.3 cm (6.24 degrees at 48.3 cm) and frequency 1.9 cycles/cm (1.6 cycles/degree at 48.3 cm) (scaled down to fit within available space). (b) ODOG predicted output intensity profile of Figure 3.2a incorrectly predicts that grey patch on the left should appear darker than the grey patch on the right. (c) Shifted White input stimulus having a patch-height of 5.3 cm (6.24 degrees at 48.3 cm) and frequency 0.74 cycles/cm (0.57 cycles/degree at 48.3 cm) (scaled down to fit within available space). At this frequency our experiments show that both grey patches should appear equally bright on average, albeit with some variance between observers. (d) ODOG predicted output intensity profile of Figure 3.2c does not corroborate with our experimental results, in this case. (e) Shifted White input stimulus

having a patch-height of 5.3 cm (6.24 degrees at 48.3 cm) and frequency 0.19 cycles/cm (0.16 cycles/degree at 48.3 cm) (scaled down to fit within available space). Observe that the illusory effect has been reversed compared to Figure 3.2a. (f) ODOG predicted output intensity profile correctly predicts that grey patch on the right should appear brighter than the grey patch on the left. (g) Predicted intensity vs. patch-height curves as obtained from ODOG filter at a frequency of 1.6 cycles/degree. (h) Experimentally observed intensity vs. patch-height curves at a frequency of 1.6 cycles/degree. (i) ODOG predicted intensity vs. frequency curves for a constant patch-height of 6.24 degrees. (j) Experimentally observed frequency characteristics for a constant patch-height of 6.24 degrees show an inversion point at 0.57 cycles/degree.

3.3 Scaling properties of Mach bands

In the previous section we saw that although the vertical length scale does not affect the perceived brightness of the grey patch in White's and Shifted-White's illusion, the horizontal length scale can significantly affect the perceived brightness of the grey patch. In this section we will see another type of scaling behavior of brightness illusions, namely the effect of length scale on the Mach band illusion [Bakshi and Ghosh, 2012a] [Bakshi and Ghosh, 2012b], which was introduced in Section 1.3.9. This effect is illustrated in Figure 3.3 where a sequence of Mach band stimuli are shown in which the size central gradient region increases stepwise from zero (Figure 3.3a) to a very large size (Figure 3.3f). So the horizontal intensity profile of first stimulus (Figure 3.3a) is a step function as plotted in Figure 3.3g (curve a), whereas the horizontal intensity profile of Figure 3.3f has the lowest slope of all the six figures. One can see that no Mach band is visible in Figure 3.3a and one can also see that as the size of the central gradient region increases, the width of the Mach Band also increases. The detailed method of measuring these changing Mach Band widths have been described in Chapter 7 of this thesis. Figure 3.4 shows an image in which every horizontal line traces a different slope of the intensity profile, depending upon its position in the vertical direction. At

the very top the horizontal intensity profile is a step function, whereas at the very bottom the horizontal intensity profile has the lowest slope. The rate of increase of the gradient region from top to bottom is constant. One can see that the width of the Mach band increases continuously from very small to very large. The Mach band region appears as if a ray of light emanating from a spot on the top and spreading out as it travels downward. It also looks as if just like a ray of light the spread of the beam is proportional to the distance travelled, which is proportional to the size of the gradient region. i.e. the width of the Mach band appears to be proportional to the width of the gradient region. This implies that the illusory effect scales in proportion to the input stimulus. This scaling behaviour implies that whatever the physiological mechanism that is generating this illusion, it can generate illusory effect of a very large range of spatial scales. This effect also cannot be explained by the ODOG filter (see Figure 5.11c in Chapter 5 of this thesis). In fact, this effect is hard to explain using traditional models of early vision because they cannot generate illusory effects of very large length scales, and will be further discussed in the concluding chapter (Chapter 7) of this thesis.

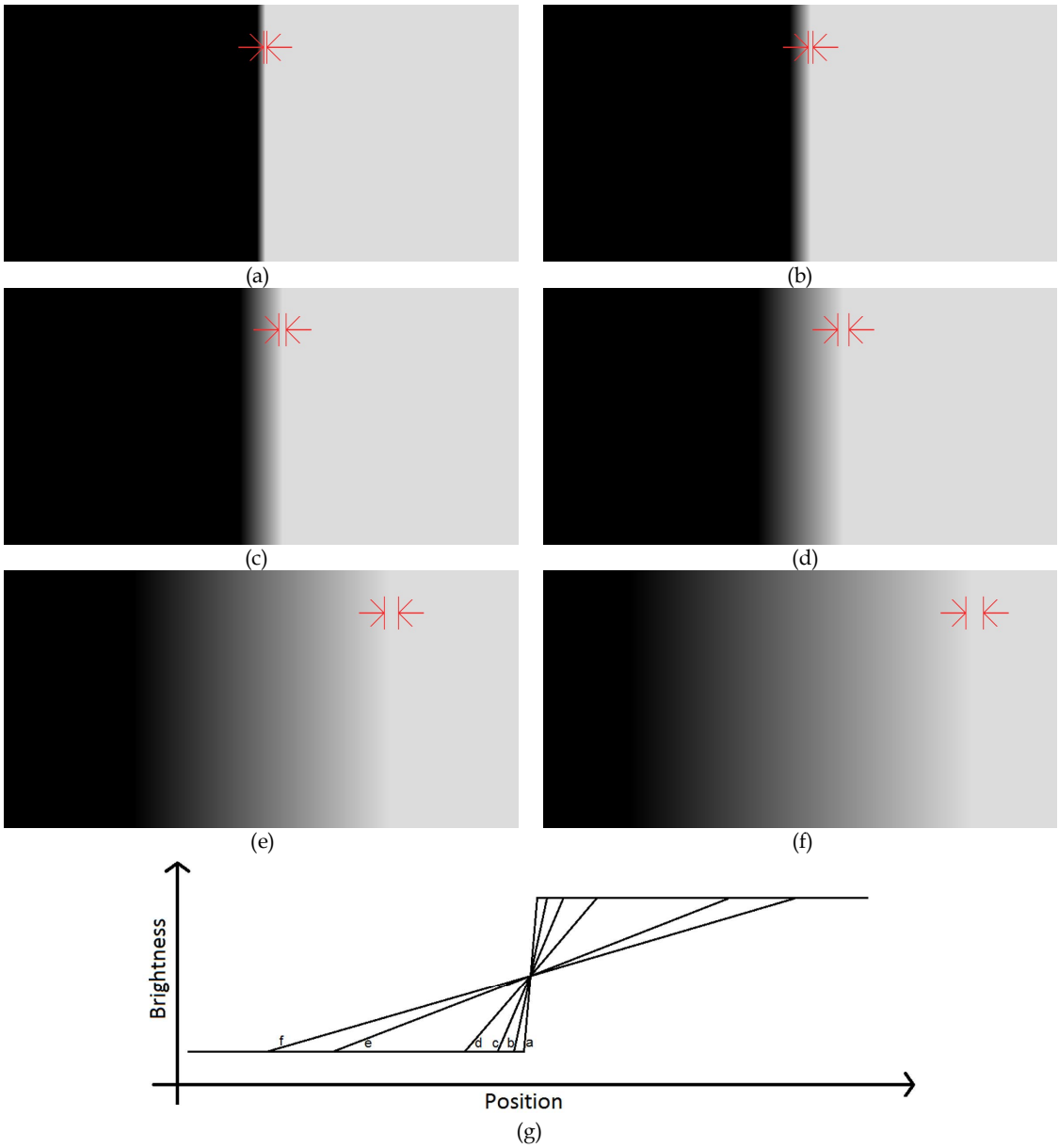


Figure 3.3: Mach band stimuli with increasing sizes of the gradient region. The widths of the gradient regions of Figures (a) through (f) are in ratio of 2:5:10:20:60:80 respectively, i.e. the horizontal gradients have been scaled down in that proportion as shown in (g).

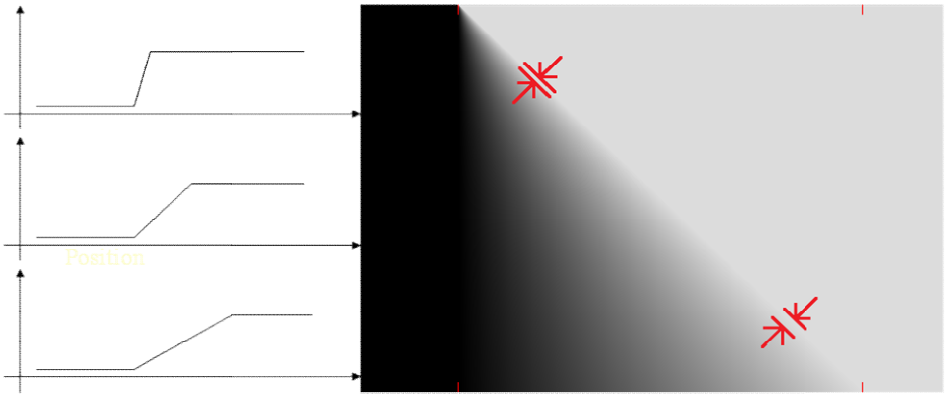


Figure 3.4: The region of intensity gradient increases linearly in size from zero at the top to its widest at the bottom. The bright Mach band is perceived thinner at the top and wider at the bottom diverging outwards like rays of light emanating from the top. The plots on the left show the brightness profiles at those corresponding levels.

3.4 Conclusions

Our experimental results establish the need for new cortical filters which not only can simultaneously explain both brightness-contrast and brightness-assimilation, but can also account for the above mentioned variations, especially, in the White and Shifted-White effects, which the ODOG filter is unable to predict.

Chapter 4

Spatial filtering based on the multi-channeled visual pathways: the Varying Contrastive Context Filter

4.1 Introduction

In this chapter we present a new model of brightness induction, that we refer to as the *Varying Contrastive Context Filter* (VCCF) model that takes inspiration from the biological structure of the human visual system in a two-fold way.

First we take inspiration from the multi-channeled structure of the visual pathways of the human visual system. Secondly we also take inspiration from the multilayered ganglion structures in the retina. Our model yields a much simpler alternative to the very well accepted ODOG model which is able to explain most brightness illusions of both brightness-contrast and brightness-assimilation types. In the next chapter we shall further develop the model in order to solve most of the problems mentioned in Chapter 3.

4.2 The multi-channeled structure of the eye-brain system

It is well known that the neural pathways that carry visual signals from the retina to the brain can be divided into at least three types, viz. *Parvocellular*, *Magnocellular* and *Koniocellular* (henceforth abbreviated as P, M, & K respectively), originating from three different types of retinal ganglion cells [De Monasterio and Gouras, 1975] [Croner and Kaplan, 1995] [Solomon et al., 2002] [Xu et al., 2001] which give rise to P, M and K channels segregated anatomically, physiologically and behaviourally [Shapley and Perry, 1986] [Silveira and Perry, 1991]. At the brain these signals arrive at a structure in the midbrain, known as the thalamus (which is often called the gateway to the cortex). Especially a substructure within the thalamus, known as the *Lateral Geniculate Nucleus* (LGN), receives most of the visual sensory signals. The LGN acts like a relay centre to the brain. From the LGN the signals are forwarded into the cerebral cortex, where most of the higher level processing, such as object recognition, is performed. The main region within the cerebral cortex that receives and processes the visual sensory data from the LGN is the *Primary Visual Cortex* (V1), located in the *Occipital Lobe* at the back of the brain. The visual cortex is also divided into two pathways, one of which is specialized for motion processing and the other for colour or form information processing [Ungerleider and Mishkin, 1982]. Several studies [Merigan and Maunsell, 1993] showed indirect evidence that M channel in sub-cortical pathways feed input to motion pathway and P channel drives

the colour or form pathway of visual cortex. The Parvocellular pathway carries much more spatial detail than the Magnocellular pathway. By selective blocking of neuronal response of either P or M channels in LGN of macaque monkeys (Macaca fascicularis and M. nemstrina), Ferrera et al. (1992) came up with results, showing that there is however, an intermixing of P and M channel contribution in the visual area V4 and many units of V1 also, providing evidence that both M and P channels probably make substantial contribution to neuronal response in colour or form pathway. It is not unlikely therefore, that all the three (including K) channels may be involved in the process of brightness perception. Additionally it has also been shown that this flow of visual signals from retina to the cortex is not strictly feedforward. It is well known that there exist significant corticothalamic feedback lines from the cortex back into the LGN [Grossberg et al., 1997] [Hupé et al., 1998]. These feedback lines can in turn modify the feedforward signals. The necessity of feedback can be clearly understood in the form of an analogy. Imagine one is listening to the radio and the volume is not sufficiently high. One then reaches out to the volume knob in order to turn up the volume. Thus one has fed a signal back into the radio, which modifies the sound coming from the radio that we wish to hear clearly, until we are satisfied with the volume level. This is a continuous process, i.e. in the next radio programme, the sound levels may change again and we may once again adjust the volume knob. This

feedback is thus capable of adjusting the relative signal strengths that are inflowing from each of the visual channels viz. P, M and K.

4.3 The ganglion layer structure of the neuronal fibres in the retina

Light rays, after being focused by the lens, fall upon the retina, which is embedded with photosensitive receptor cells that convert the light energy into electrochemical signals. These electrochemical signals after passing through a cascade of *Ganglion cells* are collected into a bundle of axonal fibres, known as the Optic fibre, which carries the signal into the brain. (In pure terms the eye could be considered an extension of the brain).

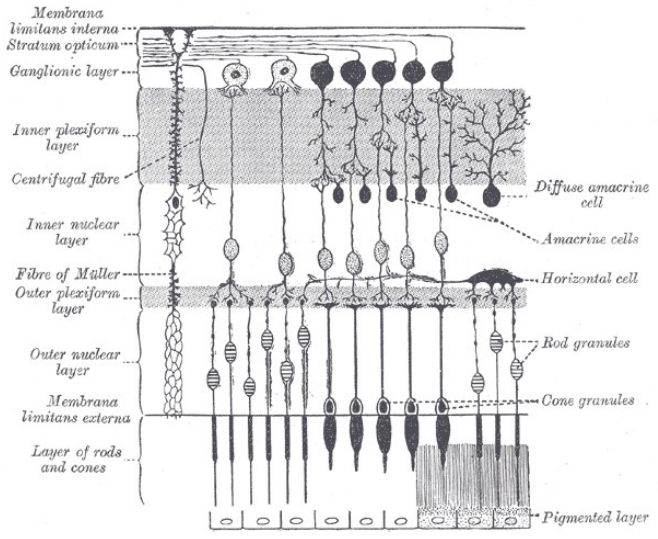


Figure 4.1: Cross section of the Retina showing the Retinal Ganglion Cells. (Public domain diagram obtained from the URL https://commons.wikimedia.org/wiki/File:Gray882.png. Original source: Henry Gray (1918) Anatomy of the Human Body.)

These Ganglion cells are capable of intermixing signals from neighbouring photoreceptor cells. By intermixing neighbouring signals they are

responsible for a phenomenon known as *Lateral Inhibition*. Lateral inhibition is the phenomenon in which a neuron's response to a stimulus is inhibited by the excitation of a neighboring neuron. Lateral inhibition has been experimentally observed in the retina and the LGN of organisms (Rodieck and Stone, 1965). Lateral inhibition makes neurons more sensitive to a spatially varying stimulus than to a spatially uniform stimulus. This is because a neuron getting stimulated by a spatially uniform stimulus is also inhibited by its surrounding neurons, thus suppressing its response. On the other hand a neuron subjected to a spatially varying stimulus is less inhibited by its neighbours that are not excited, thus producing stronger response. Therefore in the case of visual neurons, lateral inhibition makes them more sensitive to edges in the scene. Although usually described for visual neurons, lateral inhibition is also found in other sensory systems, such as auditory and olfactory neurons. The total region to which a particular neuron is sensitive is called the *receptive field* of the neuron. Therefore the receptive field of an optic nerve fiber that exits the retina may span several photosensitive rod or cone cells and will produce a distribution of responses when light falls on each of the receptor cells in its receptive field.

Lateral inhibition has been traditionally modeled using the Difference of Gaussian (DOG) function.

$$DOG(r;\sigma_1,\sigma_2) = A_1 e^{-\left(\frac{r^2}{2{\sigma_1}^2}\right)} - A_2 e^{-\left(\frac{r^2}{2{\sigma_2}^2}\right)} \quad (4.1)$$

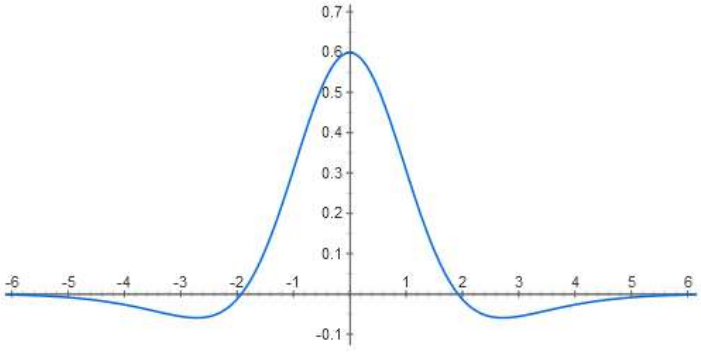


Figure 4.2: DOG response curve.

When a light spot falls directly on the center of the receptive field of a neuron it produces a large positive response but when it falls on a neighbouring neuron it produces a lower response and when light falls even further away from the centre it may produce a negative response. So the response follows a curve qualitatively similar to the DOG curve in Figure 4.2 above.

4.4 The DDOG Filter

Since a DOG like response function can be produced by the neural processing done within the retina itself, it is quite plausible that another layer of Gaussian Differencing may be performed on the DOG output signal, because as shown above, the retina consists of several layers of ganglion cells that can significantly transform the input signal. We call this additional differencing operation the *Difference of Difference of Gaussians* (DDOG). The DDOG equation may be written as:

$$\mathrm{DDOG}(\boldsymbol{r};\ \mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{A}_4, \sigma_1, \sigma_2, \sigma_3, \sigma_4) =$$

$$\mathrm{DOG}(\boldsymbol{r};\ \mathrm{A}_1, \mathrm{A}_2, \sigma_1, \sigma_2) - \mathrm{DOG}(\boldsymbol{r};\ \mathrm{A}_3, \mathrm{A}_4, \sigma_3, \sigma_4) \qquad (4.2)$$

The DDOG is actually a linear superposition of two DOG functions in opposite phases, each of which may be looked upon as representing an on-centered and an off-centered cell respectively. We use two different versions of the DDOG function each with a different set of parameter values A_1, A_2, A_3, A_4 etc. Unlike the ODOG model, our algorithm makes use of only isotropic filters so that if the input image is rotated by the certain angle then the output is guaranteed to be rotated by the same angle. Thus the simplicity of our model compared to ODOG also makes it a much more plausible model for the neural networks involved with low-level vision.

4.5 Modeling the P and M channels using the DDOG filters

The visual channels all originate in the retina and they all contain information about the entire visual field. The only way they can differ is that they must have been through different layers of ganglion processing, which makes them carry different spatial and temporal aspects of the visual signals. It is well known that the M channel has fast temporal response and low spatial resolution whereas the P channel has high spatial resolution but slow temporal response [Kandel et al., 2000] [Merigan and Maunsell, 1993].

	A_1	A_2	A_3	A_4	Sampling Interval
M-filter	10	0.5	0.5	0.08	0.5
P-filter	10	0.25	0.25	0.01	0.25

Table 4.1: Coefficient values of DDOG filter

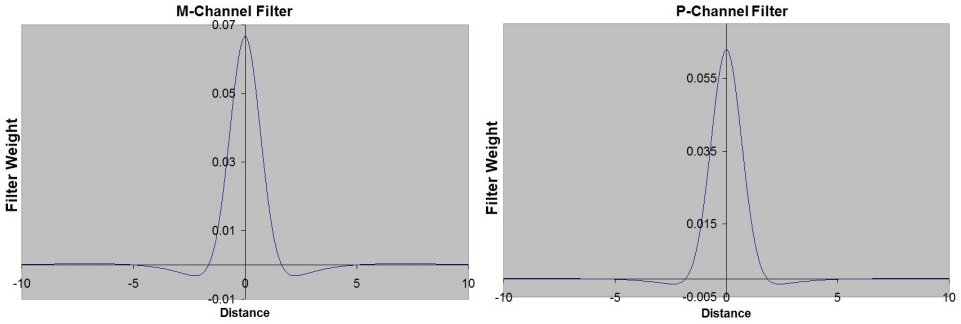
We use the values $\sigma_1=0.7$, $\sigma_2=3\sigma_1$, $\sigma_3=3\sigma_1$, $\sigma_4=9.3\sigma_1$ for both filters.

These widths are based upon the report of Shou et al. [Shou et al., 2000] that the diameter of the extended non-classical surround is at least ten times that of the classical center. The parasol/magno (M) cells which have larger size of dendritic trees, cell bodies and receptive fields, receive inputs from a relatively larger number of rods and cones [Kandel et al., 2000] and possess higher values for A_2 , A_3 , A_4 . Since the coefficients A_2 , A_3 , A_4 correspond to larger spatial scales (thereby lower spatial frequency) compared to A_1, the M-channel filter is assigned higher values for A_2 , A_3 , A_4 compared to the P-channel filter because M-channel is less responsive to smaller spatial scales compared to the P-channel. However since the sigma values are the same for both M and P-channel filters the shape and spatial extent of the filters are roughly the same and their biological dissimilarities are not reflected in their plots shown in Figure 4.3. In order to simulate the varying spatial resolutions of the different visual channels we introduce the "Sampling Interval" parameter. This

parameter is used in the discretization process of the DDOG filter which is a necessary step in the implementation of the DDOG filters. The DDOG function itself is a continuous function of the spatial variable r. But in order to be digitally implemented this function has to be sampled at a finite number of points which are separated spatially by the aforementioned Sampling Interval parameter. The M channel, which has a lower spatial resolution is assigned a higher sampling interval as specified in Table 4.1. The midget/parvo (P) cells, on the other hand, with lower spatial sampling interval and consequently higher spatial sampling frequencies, have higher spatial resolution as compared to the M cells [Merigan and Maunsell, 1993]. It should be noted here that the main goal of our model is to demonstrate the possibility of qualitatively explaining brightness-contrast and brightness-assimilation illusions simultaneously by using multi-channel filtering models. We have not attempted to model any quantitative electrophysiological experimental data of any illusory effect. Therefore the parameter values used in our model were mostly arrived through trial and error and not by using quantitative experimental data as obtained by neurophysiologists.

The M and P filters are graphically shown and compared in Figure 4.3. Notice that although the M/P filters in Figure 4.3a-4.3b roughly look just like the DOG function (Figure 4.3), there are small but crucial differences. The M/P filter-functions cut the x-axis at 4 points whereas

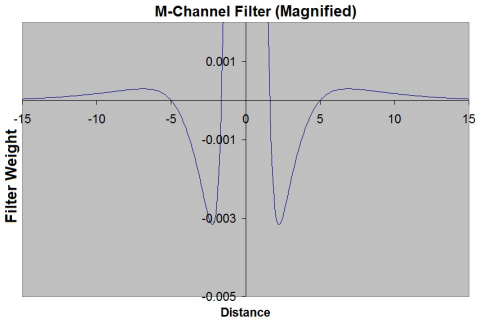
DOG cuts the x-axis only at two points. This can be clearly seen in Figure 4.3c-4.3d, which is a magnified view. From Table 4.1 it can be seen that the value of A_4 for the P filter is much smaller than the A_4 value for the M filter. For this reason, compared to the M-filter, the P-filter is very close to the DOG filter in its output. The M/P filters have five local extrema unlike DOG which has only three local extrema, although the local maxima farthest from the origin are so small they are not visible in Figure 4.3a-4.3b. Figures 4.3g-4.3i show the frequency response curves of M & P channels. Even though the frequency response curves are largely that of a low pass filter, but the frequency response of M-channel shows an extra depression at very low frequencies. This difference turns out to be important for DDOG model to work properly. Figure 4.3f demonstrates that there is a clearly visible disinhibition, i.e. surround suppression in the M i.e. parasol channel, as compared to the P, i.e. midget channel. This, in fact, is the reason why the M channel shows an extra depression in the frequency response curve in Figure 4.3i. Solomon et al. [Solomon et al., 2002] have experimentally substantiated the fact that extra-classical surround suppression which was found in parasol ganglion cells was virtually absent in the corresponding midget cells.


M-Channel Filter
Filter Weight
Distance
0.07
0.05
0.03
0.01
-0.01
-10
-5
0
5
10

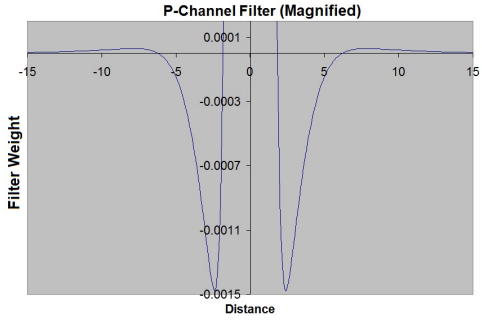
(a)

P-Channel Filter
Filter Weight
Distance
0.055
0.035
0.015
-0.005
-10
-5
0
5
10

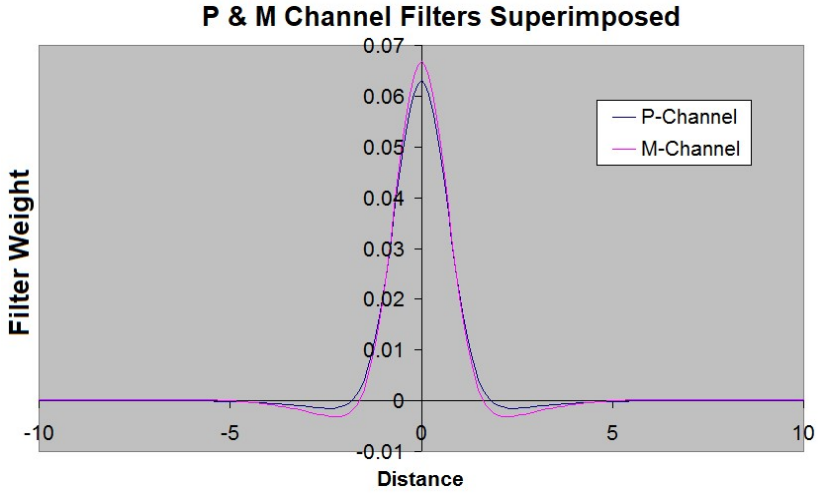
(b)


M-Channel Filter (Magnified)
Filter Weight
Distance
0.001
-0.001
-0.003
-0.005
-15
-10
-5
0
5
10
15

(c)


P-Channel Filter (Magnified)
Filter Weight
Distance
0.0001
-0.0003
-0.0007
-0.0011
-0.0015
-15
-10
-5
0
5
10
15

(d)


P & M Channel Filters Superimposed
Filter Weight
Distance
P-Channel
M-Channel
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
-10
-5
0
5
10

(e)

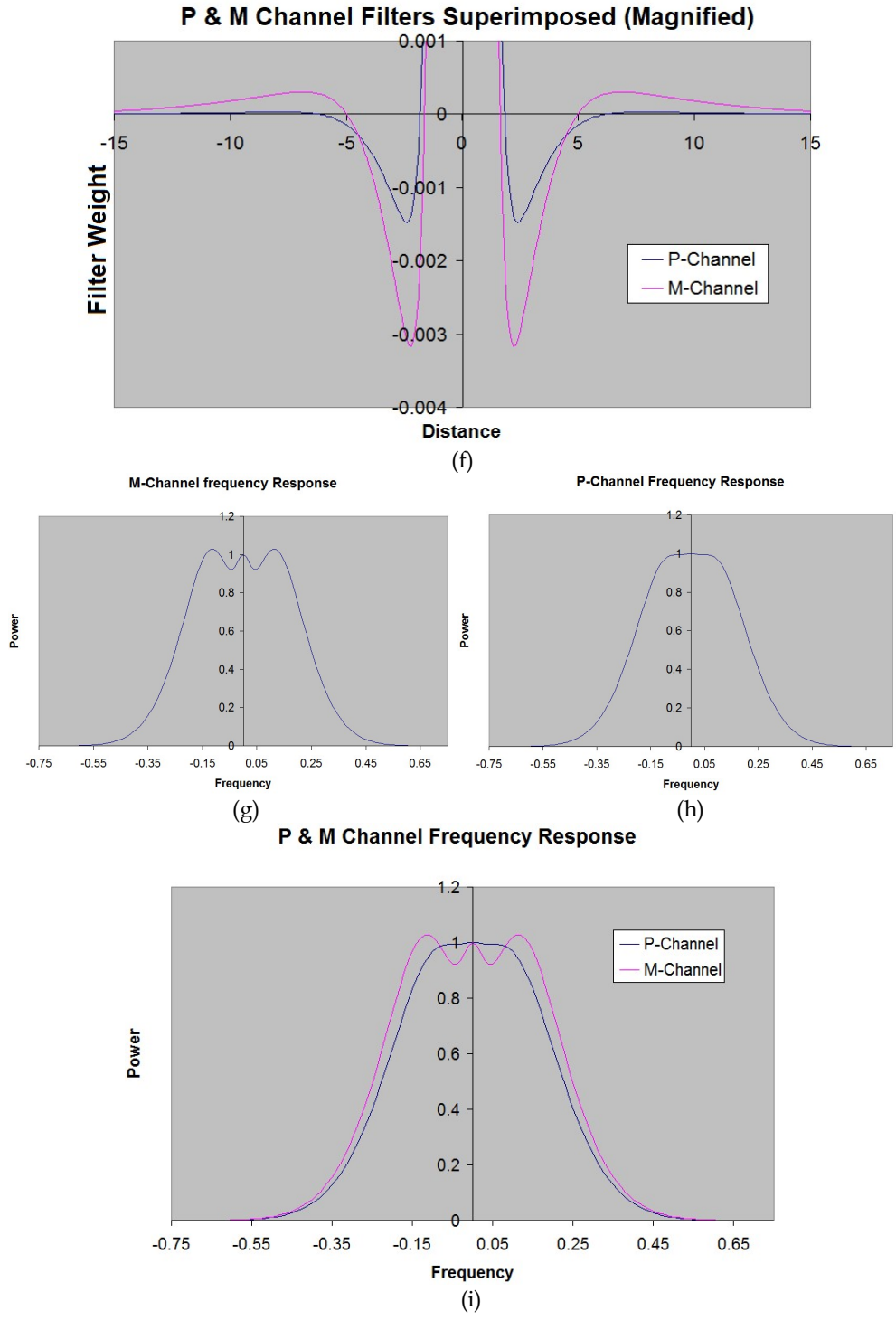


Figure 4.3: Graphs of P & M channel filter functions. (a) Graph of M-channel filter function with respect to distance from the filter center. (b) Graph of P-channel filter function. (c) Magnified view of M-channel filter showing that it cuts the x-axis at four points (unlike DOG which only cuts the x-axis at two points). (d) Magnified view of P-channel filter, which like the M-channel filter also crosses the x-axis a four points, but

after crossing back into the upper side of the x-axis it rapidly settles to zero. Therefore the P-channel is very close to the DOG filter.

We propose that Equation (4.2) with different parameter values (A1, A2, A3, A4, σ1, σ2, σ3, σ4, and the sampling interval of Table 4.1) can mimic the role of both Parvocellular and Magnocellular pathways, thus modeling the two major complementary channels in the central visual pathway [Merigan and Maunsell, 1993]. Since the M channel has lower spatial resolution than the P channel, we have used a lower sampling resolution to implement the M channel than that used for the P channel. This is an important point of dissimilarity between the response produced by the M and P channel filters which are not reflected in Figure 4.3.

4.6 The VCCF model

We now propose the *Varying Contrastive Context Filter* (VCCF) model based upon the principle that the final brightness percept is formed in the visual cortex through a linear combination of the M and P channels [Bakshi et al., 2021]. For the sake of simplicity, we leave out the Koniocellular channel, whose role is not so well defined as in the case of the other two channels, and we base the proposed VCCF model upon only two spatial filters representing the Magnocellular and Parvocellular channels. The degree to which the delayed P output is combined with the initial M output, depends upon a concept that we term as the *contrastive context* of the incoming stimulus. This we model by introducing a

parameter α, which we refer to as the *Factor of Contrastive Context* (FOCC). Qualitatively, the greater the number of sharp edges in a stimulus the lower the value of the FOCC, and vice versa. The more the role of such contrastive context, the greater will be the component of P alongside the initial M output in forming the brightness percept. So in our proposed model, the outputs of the M channel and the P channel are finally combined through the following equation:

$$VCCF(\alpha) = \alpha P + (1 - \alpha)M \quad (4.3)$$

where P and M correspond to their respective DDOG expressions (4.2) and α is a weight value representing the Factor of Contrastive Context. We are going to show that α, that varies between 0 and 1, is either very high (close to 1) or medium (close to 0.5), according as whether the illusion type is brightness contrast (in which case understandably, the FOCC is high because of a low number of edges) or brightness assimilation (in which case the FOCC is low because of high number of edges). This implies using Equation (4.3) that in case of brightness contrast, contextual vision through P plays a major role along with a smaller contribution from M, unlike brightness assimilation when P plays an almost competing role with M to produce the final brightness percept. Such a theory corroborates with the prevalent view [Blakeslee and McCourt, 2004] that when the stimulus mainly loses its low frequency content due to spatial filtering by high spatial frequency tuned channels, which is P as per our assumption

[Ghosh, 2012], then brightness contrast appears; on the other hand if the low spatial frequency tuned channels, i.e. M according to our assumption [Ghosh, 2012], mainly filter out the high frequency content of the stimulus, when value of α comes down to 0.5 or less, it results in brightness assimilation. Of course, in the second case, the major contrast edges cannot be wiped out; otherwise the image will lose its key information. This is why in this case, P must also have a sufficient share in the computation of the final output, rather than coming down to a value close to zero so that the brightness percept depends solely upon M. The VCCF filter coefficients used to implement the P channel and M channel are shown in Table 4.1.

4.7 Sample VCCF results

Since there is no neurophysiological data to model/quantify the contrastive context (CC) to the best of our knowledge, hence in order to experimentally obtain suitable values of alpha (α) in Eqn. (4.3), in this section we apply the VCC filter to certain sample illusions and analyse the results to arrive at the values of α which are appropriate for brightness contrast (i.e. high FOCC) or brightness assimilation (i.e. low FOCC). The approach is like that of curve-fitting, where we first pass the stimuli through the VCC filter for various values of α. From the resulting output we find a minimum possible set of values of α for which the output matches human observation. From the resultant alpha ranges for various

stimuli we empirically select two values of α, one close to 0.5, which is α = 0.45 for the low FOCC stimuli, i.e. the brightness assimilation illusions and another double of that value, close to unity, which is α = 0.9 for the high FOCC stimuli, i.e. the brightness contrast illusions. These are two values at which the VCC filter can successfully model the illusory effect, although this is true within a range around each of these values. All the subsequent stimulus images are scaled down versions of the actual stimuli used in our experiments.

4.7.1 Sinusoidal grating

The sinusoidal grating illusion, as shown in Figure 4.4a, consists of a spatial field of sinusoidally varying background intensity over which a thin gray strip of uniform intensity is placed. Despite having a uniform intensity, the strip seems to have an undulating brightness variation along its length. The perceived brightness of the gray strip changes in a direction opposite to that of the background intensity i.e., in places where the background is dark, the gray strip looks bright and vice versa. Since the background changes smoothly, and the test patch in foreground is of uniform luminance, it follows that the number of contrast edges in this stimulus is very low. Hence this is a type of brightness contrast illusion (see section 2.2.3) with high FOCC.

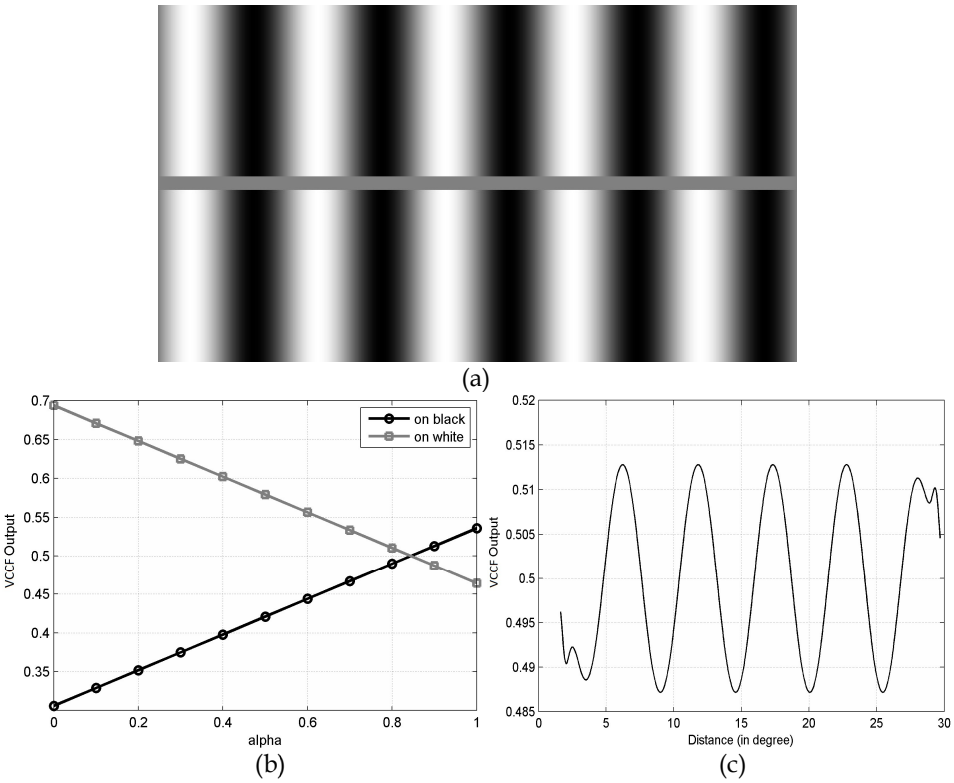


Figure 4.4: (a) Sinusoidal grating stimulus with spatial frequency of 0.18 cycles/degree and gray strip width of 0.56 degrees. (b) Intensity output versus alpha graph from VCCF model for input stimulus as shown in (a) (c) VCCF predicted output intensity profile for grating stimulus in (a) at alpha = 0.90.

We have run the VCC filter for various values of alpha and plotted the predicted output brightnesses (i.e. the minimum and maximum brightness values over the central gray strip) with respect to alpha in Figure 4.4b. From Figure 4.4b one can see that for alpha values in the range of [0.85, 1], the VCC filter gives the desired response, i.e., the predicted brightness of the gray strip over a background maxima is lower than the predicted brightness over a background minima. Figure 4.4c shows the VCCF predicted brightness profile of the gray strip at the chosen alpha value of 0.9.

4.7.2 Simultaneous brightness contrast (SBC)

The Simultaneous Brightness Contrast (SBC) illusion (Figure 1.2a) consists of two identical gray patches surrounded by two different intensity backgrounds. One grey patch lies within a bright surrounding while the other lies in dark surrounding. The gray patch in the dark surrounding seems brighter than its counterpart. Hence this is also a type of brightness contrast illusion having mostly uniform luminance values and less number of edges and consequently with high FOCC.

We have run the VCC filter on the SBC stimulus for different values of alpha and plotted the predicted output brightness values of the two gray patches with respect to alpha in Figure 4.5b. From Figure 4.5b it can be seen that for alpha values in the range of [0.65, 1] the VCC filter gives the desired response, i.e., the predicted brightness of the gray patch surrounded by white is lower than the predicted brightness of the gray patch surrounded by black. So for both the sinusoidal grating as well as the SBC, 0.9 is a value within the range of the desired psychophysical experience, since $0.65 < 0.85 < 0.9 < 1$. Figure 4.5c shows the VCCF predicted brightness profile calculated along the horizontal centerline of the SBC stimulus at the chosen alpha value of 0.9.

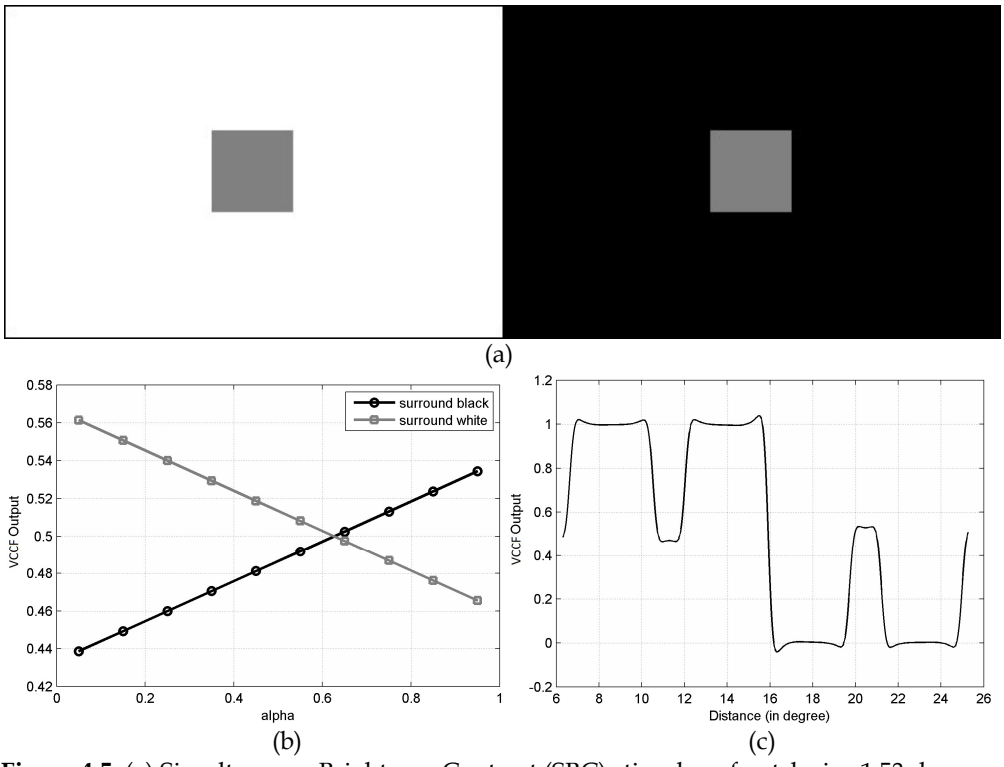


Figure 4.5: (a) Simultaneous Brightness Contrast (SBC) stimulus of patch size 1.53 degrees (b) Intensity output versus alpha graph from VCCF model for input stimulus as shown in Figure 4.5a (c) VCCF predicted output intensity profile for SBC stimulus in Figure 4.5a at alpha = 0.90.

4.7.3 White's illusion

White's stimulus consists of a vertical grid of alternating black and white bars, upon which two gray patches of identical intensities are placed, one over a white bar and the other over a black bar (Figure 4.6a). From Figure 4.6a it can be seen that the gray patch lying on the coaxial white bar (on the right) seems apparently darker than the gray patch lying over the coaxial black bar (on the left). Unlike the case of SBC above, this is a type of brightness assimilation illusion (see Section 2.2.3) with many edges and low FOCC. As described in the previous sample illusions, we have plotted the VCCF predicted output brightness values of the two gray patches with

respect to alpha in Figure 4.6b, from where it can be seen that for alpha values in the range of [0, 0.65] the VCC filter yields a response consistent with brightness assimilation. Figure 4.6c shows the VCCF predicted horizontal brightness profile of White's stimulus at our selected alpha value of 0.45 which is half the value we used for the high FOCC stimuli above.

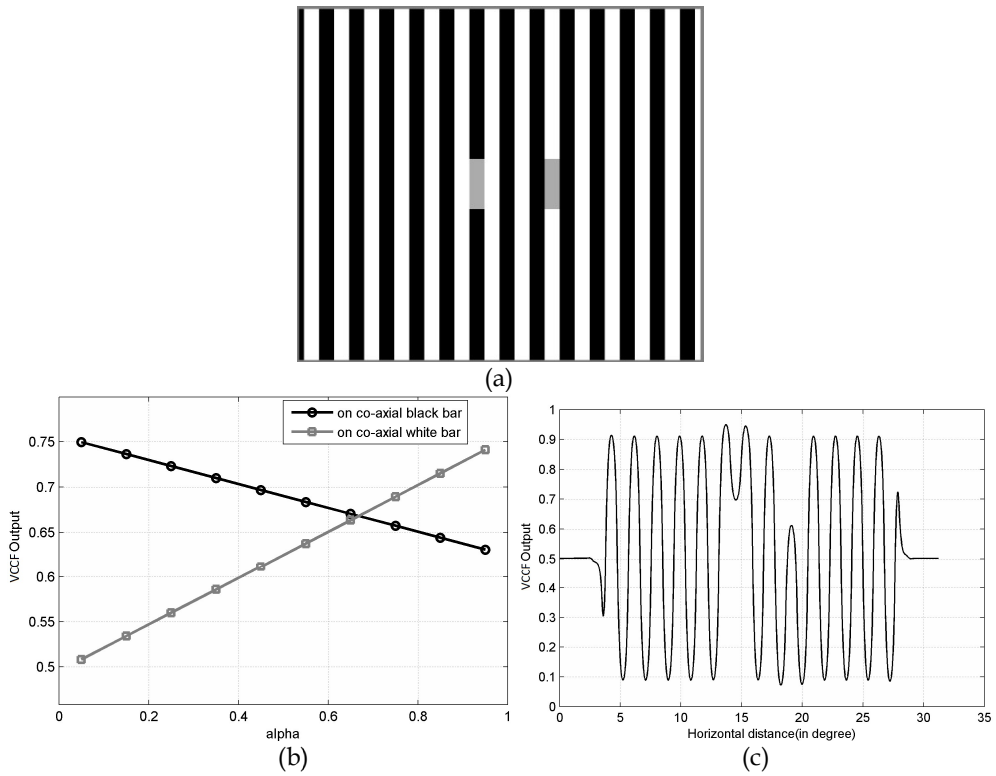


Figure 4.6: (a) White's illusion with frequency of 0.53 cycles/degree and patch height of 3.12 degrees. (b) Intensity output versus alpha graph from VCCF model for the input stimulus shown in Figure 4.6a. (c) Intensity profile given by VCCF at alpha value of 0.45.

4.7.4 Shifted White's illusion

The Shifted White's illusion (Figure 4.7a) is a modification of White's stimulus wherein a horizontal section containing the gray patches is shifted sideways by one barwidth. Just like White's effect, this illusion also

is a type of brightness assimilation effect with a lot of edges and low FOCC. Figure 4.7b shows a plot of the VCCF predicted brightness values of the two gray patches for various alpha values. It can be seen that $\alpha \in [0, 0.65]$ reproduces the desired brightness assimilation effect. Figure 4.7c shows the VCCF predicted horizontal brightness profile of the Shifted White's stimulus at our selected alpha value of 0.45.

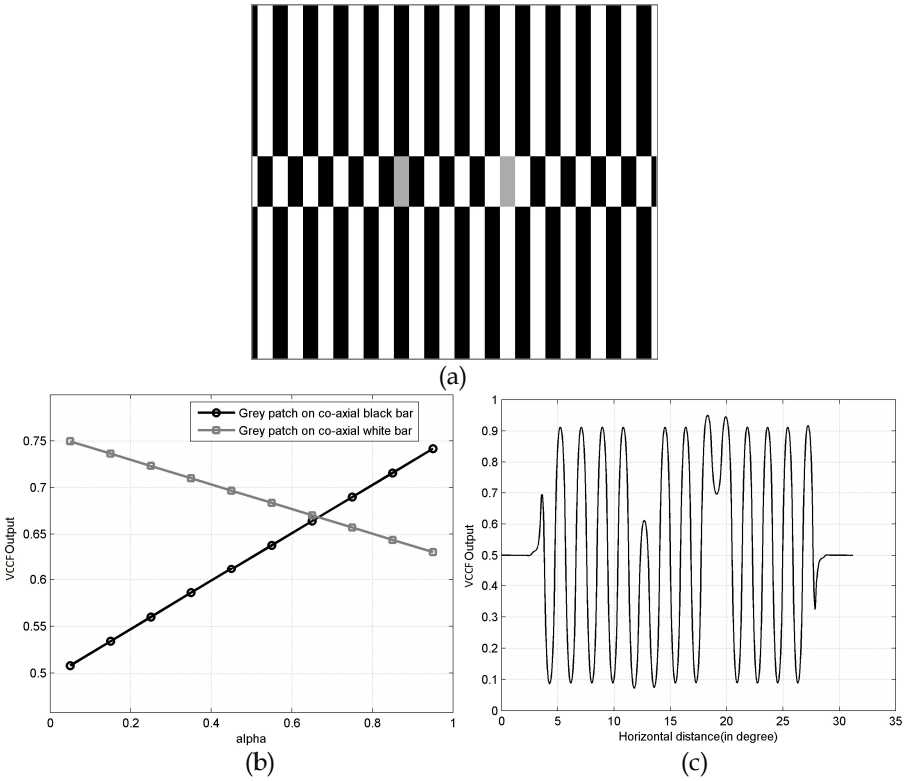


Figure 4.7: (a) The Shifted White's illusion with frequency of 0.53 cycles/degree having patch height 3.12 degrees. (b) Intensity output versus alpha graph from VCCF model for input stimulus as shown in 4a.(c) Intensity profile given by VCCF at alpha value of 0.45.

4.8 Comparison of VCCF results with ODOG in terms of extended patch-height in brightness assimilation illusions

In section 4.7 we presented four visual illusions, two of brightness-contrast (i.e. high FOCC) and two of brightness assimilation (i.e. low FOCC) types, and in each of those cases the VCCF model was successfully able to explain the illusion. However, these cases can also be accounted for by the ODOG model [Blakeslee and McCourt, 1999] [Blakeslee and McCourt, 2004]. The illusions shown in sections 4.7.3 and 4.7.4 contained gray patches of relatively moderate height (with respect to the stimulus height). As discussed in Section 3.2, when the gray patch height is extended to relatively large values the ODOG filter fails to predict the illusory effect. However the proposed VCCF based model continues to correctly explain the perceived illusion. Below we present a comparison between ODOG and VCCF models. We first show the predicted brightness profiles of the two models at extended patch lengths. Then we show plots of the Intensity versus Patch-height Characteristic Curve (IPCC) of both the models. Finally we compare these IPCC curves with the experimentally obtained IPCC curve, to verify their validity.

4.8.1 White's illusion

White's illusion, as shown in Figure 4.8a, belongs to the class of brightness assimilation illusions. The gray patches shown in Figure 4.8a have a large patch-height of 9.35 degrees. The ODOG predicted brightness profile of Figure 4.8b clearly shows the gray patch on the left to be brighter than the one on the right, which is incorrect as can be verified by direct

observation. Our proposed model predicts the correct direction of brightness change as shown in Figure 4.8c.

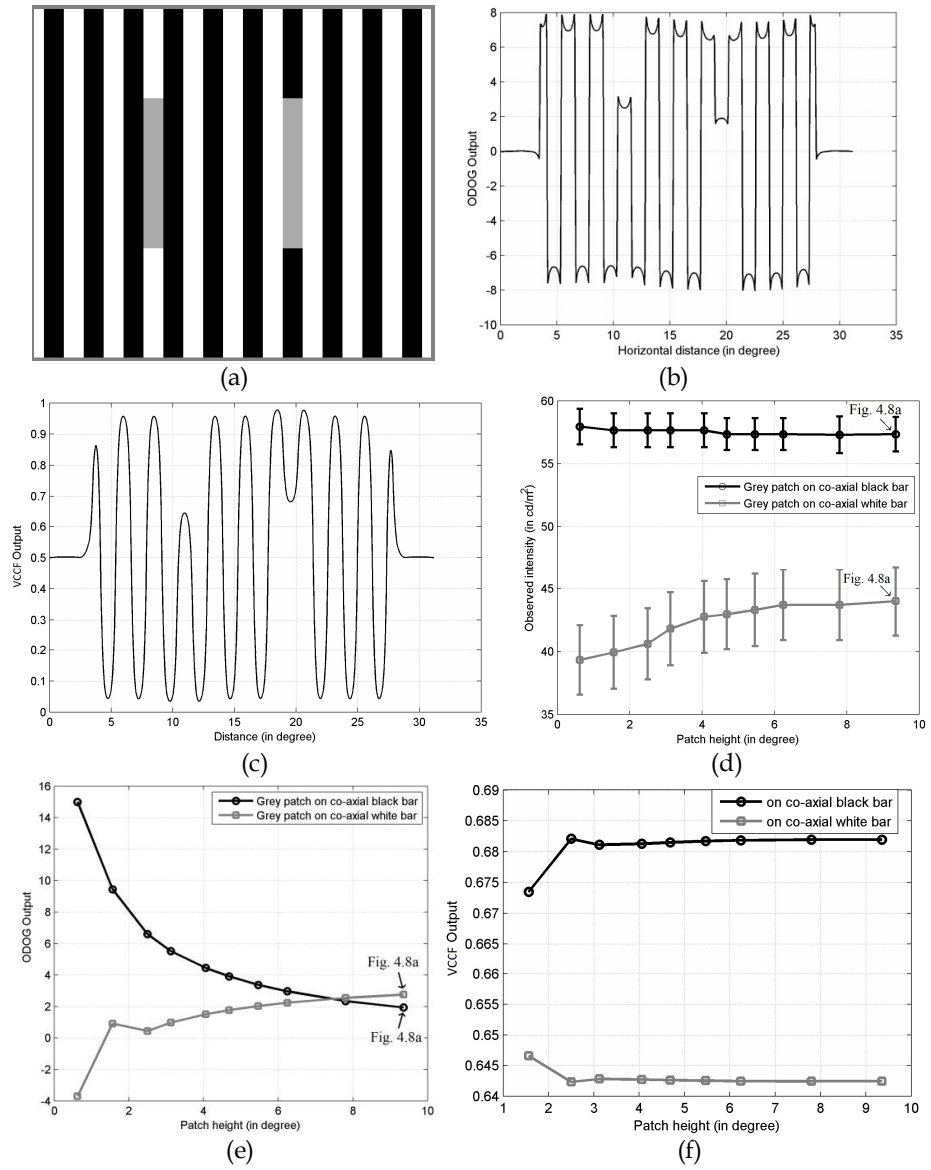


Figure 4.8: (a) Input stimulus of Whites illusion with patch height of 9.35 degrees and frequency 0.40 cycles/degree (b) ODOG output intensity profile for stimulus shown in (a) which fails to predict the direction of brightness change of White's effect. (c) Output intensity profile from VCCF model for stimulus shown in (a) which correctly predicts the direction of brightness change of White effect. (d) Experimental results of subject's observation of intensity versus patch height at a constant frequency of 0.40 cycles/degree. (e) ODOG predicted intensity versus patch height characteristic graph at a constant frequency of 0.40 cycles/degree. When patch height goes over 7.48 degree, ODOG model shows an inversion in its perceived intensity prediction which clearly contradicts human observations as shown in (d). (f) VCCF predicted intensity versus patch height

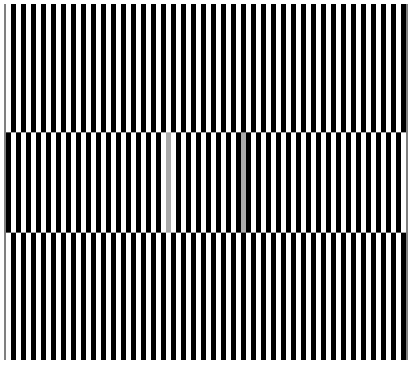
characteristic graph at a constant frequency of 0.40 cycles/degree. VCCF model shows no inversion in its prediction and accurately matches the direction of brightness change of White's effect as indicated by the experimental graph in (d).

Figures 4.8a-4.8c illustrate one specific instance in which the proposed VCCF based model correctly predicts the direction of brightness change for which the ODOG filter fails. In order to further investigate the range of patch-lengths for which this trend continues, we varied the gray patch height over a wide range of values from 0.63 degrees to 9.35 degrees while recording the VCCF and ODOG predicted outputs, and compared them with experimental observations. The Intensity versus Patch-height Characteristic Curves (IPCC) for the experimentally observed, the ODOG predicted and the VCCF predicted values are respectively plotted in Figures 4.8d-4.8f. The ODOG predicted IPCC shows an intersection point at 7.48 degrees patch-height, where a reversal of the direction of brightness change occurs. We refer to this particular point of inversion as the *Threshold Inversion Point* (TIP). The VCCF predicted IPCC shows no such inversion just like the experimentally observed IPCC.

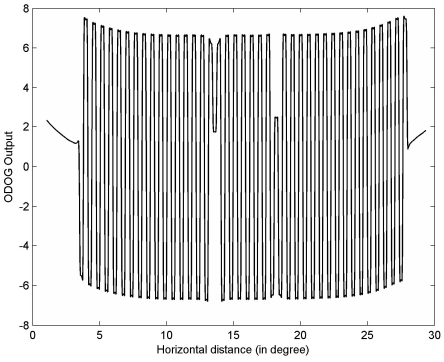
4.8.2 Shifted White's illusion

Figure 4.9a shows the Shifted White's illusion which, like White's effect, also belongs to the brightness-assimilation class of illusions. The gray patches have a patch-height of 6.24 degrees. The ODOG profile as shown in Figure 4.9b predicts the left-hand patch to be darker than the right-hand patch, which is incorrect, as can be directly seen from Figure 4.9a. Figure

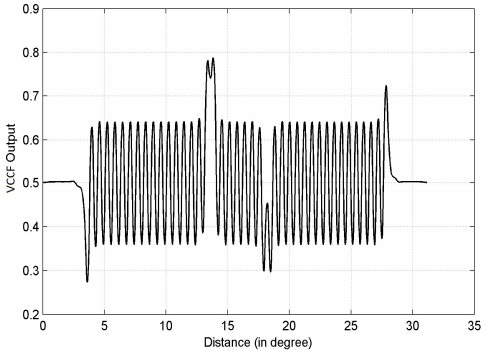
4.9c shows the corresponding VCCF predicted brightness profile, which clearly shows the left-hand patch to be brighter than the right one. The corresponding IPCC plots for the Shifted White's illusion (with patch-height ranging from 0.75 degrees to 9.35 degrees) are shown in Figures 4.9d-4.9f. As for the case of Whites illusion above, the ODOG predicted IPCC shows an inversion of brightness values beyond a *Threshold Inversion Point* (TIP); (TIP= 4.8 degrees here). The VCCF predicted IPCC shows no inversion, which agrees with the experimental IPCC.



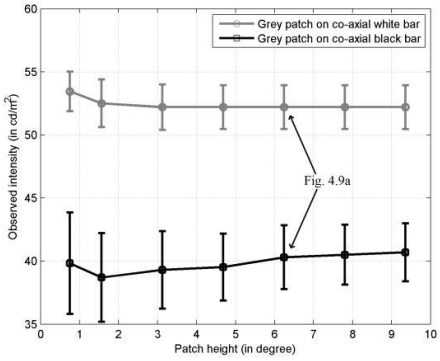
(a)



(b)



(c)



(d)

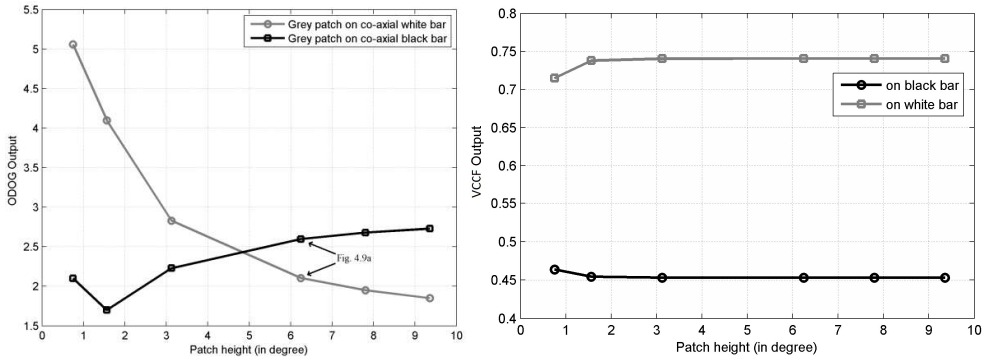


Figure 4.9: (a) Shifted White's stimulus with test patch height of 6.24 degrees and frequency 1.6 cycles/degree. (b) ODOG predicted output intensity profile for stimulus shown in (a) which incorrectly predicts the direction of brightness change of the Shifted White's effect. (c) Output intensity profile from VCCF model for stimulus shown in (a) which successfully predicts the direction of brightness change in the Shifted White's effect. (d) Experimental subject observation of intensity versus patch height at a constant frequency of 1.6 cycles/degree. (e) ODOG predicted intensity versus patch height characteristic graph at a constant frequency of 1.6 cycles/degree. The ODOG model shows an inversion in the perceived intensity prediction when patch height is above 4.8 degree, which goes against the experimental observations shown in (d). (f) VCCF predicted intensity versus patch height characteristic graph at a constant frequency of 1.6 cycles/degree. The VCCF curve shows no inversion in its prediction and agrees with the experimental graph in (d).

4.9 Conclusion

We have proposed a Varying Contrastive Context Filter (VCCF) based model of brightness perception. Our approach is based on the DDOG filter, which is an extension of the pre-existing Difference of Gaussians (DOG) model by introducing another layer of differencing leading to the Difference of Difference of Gaussians (DDOG). Our proposed model is further inspired by the fact that the human visual system combines the incoming signals arriving via the Parvocellular and Magnocellular pathways at the LGN. In similar fashion we too mix the outputs of two separate M and P spatial filters distinguished by different values of their

defining parameters. Although the filters may be combined in numerous possible ways, we have chosen the simplest possible strategy, i.e., a linear combination of the two filters. The weight factor, alpha (α), used for the linear combination is termed as the Factor of Contrastive Context (FOCC) in our model. By analyzing various brightness-contrast as well as brightness-assimilation illusions we arrive at a minimal set of values of alpha, which are able to successfully predict the direction of brightness shift in brightness contrast (i.e. high FOCC) as well as the brightness assimilation (i.e. low FOCC) perception. We compared the proposed model against the already well established ODOG filter. Extensive tests and simulations suggest that for most illusions both ODOG and VCCF produce correct output, but for certain cases in which the ODOG filter fails to predict the illusory effect, our proposed VCCF model continues to be effective. We have also experimentally validated our results with observations from human test subjects. The fact that the FOCC assumes one set of values out of the aforesaid minimal set of values in the case of brightness contrast illusions and another distinct set of values in the case of brightness assimilation illusions, indicates the important role of feedback in the human visual system. This feedback necessarily has to arrive from the upper stages of the visual pathway and modulates the lower stages of the pathway according to the contrastive context. We hypothesize that this feedback likely originates from the cortex and

controls the amount of mixing of the P and M channels in the LGN through the cortico-geniculate synapses. Thus the VCCF model emphasizes the importance of considering contextual feedback as an important variable in models of visual perception. Although the VCCF model has one free parameter, it is still much simpler than the ODOG model. ODOG uses oriented difference of Gaussian filters at 7 different length scales and 6 different orientations (i.e. a total 42 ODOGs) and non-linearly combines them. The length scales and weighting factors, although fixed, are far greater in number than the VCCF model. The VCCF model, in effect, uses 8 isotropic Gaussians with 4 length parameters and 8 weight factors. The free variable alpha, which from the observations made in this chapter, seems to depend on the number of edges in the stimulus, ideally must be computed from the stimulus itself, and in the next chapter we figure out what may determine the contextual factor alpha. However, we still observe that a binary choice of one out of two values of alpha can explain most of the brightness perceptions. This indicates that just a small amount of feedback from the upper stages is enough to determine the amount of mixing of the P and M channels.

Chapter 5

The DDOG based Adaptive Contrastive Context Filter

5.1 Introduction

In the previous chapter we proposed the DDOG based VCCF model in which we observed how several of the complexities of brightness perception may be explained by a mixture of the P and M channels which mix in a proportion determined by feedback from the upper stages of the visual system, which we call the Factor of Contrastive Context (FOCC). But we left the mixing parameter (alpha) free, observing only that a minimal set of values of the mixing parameter could correctly predict the direction of brightness shift in various types of illusions, and the fact that its value depends upon the number of edges in the stimulus. In this chapter we make an attempt to automate the process of determining the Factor of Contrastive Context (FOCC) through a two-pass model of multi-channeled vision, which we term as the Adaptive Contrastive Context Filter.

5.2 The two-pass model for multi-channeled vision

As mentioned in the previous chapter, the signals from the eyes enter the brain via several channels, namely, the Parvocellular, Magnocellular and Koniocellular channels, each of which carries information about the entire visual field. We also hypothesized how the information from all these channels may be fused by the brain to produce a single final percept and how brightness illusions could be explained by the merging of these independent channels. To explore this idea we propose the two-pass model of multi-channeled vision. Our model, as in the case of VCCF, uses only the two major channels, i.e. the P-channel and the M-channel. This is done for simplicity, since we are only trying to explore step-by-step if this way of thinking can explain all the various types of brightness illusions.

It has been known for a long time that signals are conducted faster through the Magnocellular channels compared to the other channels [Kaplan and Shapley, 1982] [Schiller and Malpeli, 1978]. Activity transferred through the Magnocellular neurons of the LGN reach area V1 some 20 ms earlier than the activity transferred through the Parvocellular neurons of the LGN [Nowak et al., 1995], which shows that despite the two channels converging beyond layer 4C, M activity precedes P activity in the different layers of V1. Based on latencies of visual responses of neurons [Maunsell et al., 1990] in different cortical areas, Bullier [Bullier, 2001] argues that such characteristics of the M channel like high contrast

sensitivity, poor chromatic selectivity, larger receptive fields and lower spatial resolution are well suited for a first-pass *'vision at a glance'*. So, although the Parvocellular pathway carries much more detail, owing to the higher spatial resolution of the midget cells of this channel, compared to the Magnocellular pathway, yet the Magnocellular pathway can carry an overall holistic information much faster than the Parvocellular pathway [Kandel et al., 2000] [Merigan and Maunsell, 1993].

According to our proposed 2-pass model of attentive vision the visual process is divided into two stages. In the first stage, called '*vision at a glance'*, the brain first interprets the contents of the Magnocellular pathway. If it can find sufficient detail in this stage itself, then it almost ignores the contents of the Parvocellular pathway. In other words if the brain can obtain sufficient information content about its environment from this channel alone then it minimally interprets the other channel. If it cannot find sufficient detail then it enters the second stage, which is called '*vision with scrutiny*', in this stage the brain examines the contents of the Parvocellular pathway to find further details in those regions where sufficient details were not found. This two stage process can be described as follows:

If

(the initial M channel output identifies that the background around the test patch is uniform)

Then

P-channel is invoked and the brightness percept is formed by P.

Else

M-channel produces the percept.

The condition of background uniformity under the *If* condition can be implemented in a variety of ways. In the proposed ACCF model described below, we, inspired by the observations made in the previous chapter pertaining to the dependence of FOCC on the number of edges in the stimulus, stick to a very simple method of evaluating background uniformity using the Laplacian operator. The Laplacian is the lowest order isotropic derivative (two-dimensional in the context of image) suitable for detecting edges in all directions in an image. It is also computationally simple since it is a sum of two operators along the x-axis and y-axis respectively. But it is possible that some other method of judging background uniformity may yield even more accurate results in the ACCF approach.

5.3 The ACCF model

The ACCF model works by adaptively combining the M and P filters, as introduced in the VCCF model in the previous chapter, in a proportion

which itself is a function of the input image. In this way the ACCF algorithm is effectively nonlinear in nature. With this model we have been able to explain both SBC and White effect and also Howe's [Howe, 2001] smooth transition between the two. In our model we implement the Magnocellular and Parvocellular channels using the same M and P filters of the VCCF model as summarized in Table 4.1 of the previous chapter. The ACCF algorithm therefore also uses isotropic filters, so that if the input is rotated by some angle then the output is guaranteed to be rotated by the same angle. The simplicity of this model thus makes it a plausible model for the neural networks involved in low and mid level vision.

The method of combination of the above mentioned two sets of M & P filters is inspired by the aforementioned 2-pass model of vision according to which the visual process is divided into two stages. In the first stage, called vision at a glance, the brain first interprets the contents of the Magnocellular pathway. The Magnocellular pathway, although lower in resolution, can carry information much faster than the Parvocellular pathway and therefore it is processed in the first stage of the process. If it can find sufficient detail in this stage then it ignores the contents of the Parvocellular pathway. If it cannot find sufficient detail then it enters the second stage which is called vision with scrutiny. In this stage the brain examines the contents of the Parvocellular pathway to find further details in those regions where sufficient details were not found.

So the above mentioned two stage process can be described as follows:

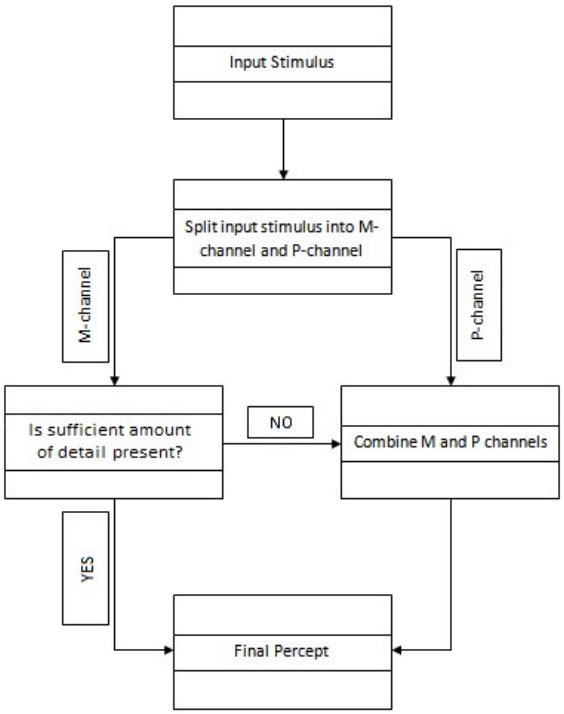


Figure 5.1: Flowchart showing the 2-pass model of attentive vision which motivates the ACCF model.

In our method, we first calculate the Laplacian (one may also remember the Laplacian of Gaussian operator in the light of David Marr [Marr, 1982] in this context) at every pixel of the M channel output image and take its square value. This gives us a positive number at every pixel. We then compute its average value per pixel. This gives us a single positive number λ for every image. If this number λ is very low then the background must be very uniform and therefore the output must be determined by the P channel. On the other hand if λ is large then the background must be highly non-uniform and therefore the output must be formed by the M channel. So for some threshold value of λ we have to

switch from P to M channel. Instead of using such a threshold value we choose to have a more gradual transition from P to M by linearly combining both the M output and the P output in some proportion *f(λ)* depending on λ. *f(λ)* must be such that, if *λ* is small the proportion of M is small, consequently *f(λ)* should be close to 0; whereas if *λ* is high then the proportion of M is high, so *f(λ)* should be close to 1. Although there exist many such functions, we chose the following form for *f(λ)*:

$f(\lambda) = \lambda / (\lambda + \lambda_0)$

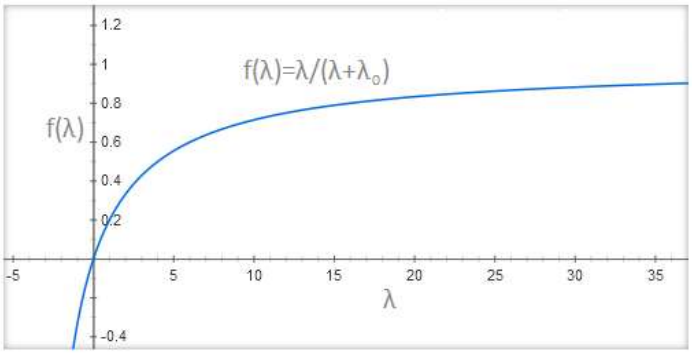


Figure 5.2: *f(λ)* vs. *λ*

Therefore the final output is determined by the following equation:

*FINAL OUTPUT = f(λ) * (M CHANNEL OUTPUT) + [1 - f(λ)] * (P CHANNEL OUTPUT)*

5.4 Implementation specifics of the ACCF algorithm

The following steps show how the ACCF algorithm has been implemented:

- Step 1: Read input Image *I*.
- Step 2: *M=M_filter(I)*. (Apply M-filter on I)

- Step 3: $L=\nabla^2(M)$. (At each pixel of M calculate Laplacian to determine level of non-uniformity)
- Step 4: $\lambda=Mean(L^2)$. (Compute mean square value of L per pixel. λ is a measure of the non-uniformity of the entire image)
- Step 5: *Compute* $f(\lambda) = \lambda / (\lambda + \lambda_0)$. (This step maps $\lambda \in [0,\infty]$ to $f(\lambda) \in [0,1]$)
- Step 6: $P=P\_filter(I)$. (Apply P-filter on I)
- Step 7: $F= f(\lambda)*M+[1- f(\lambda)]*P$. (Compute final output F)

Steps 3 & 4 are the analogues for the process of finding sufficient details in the input stimulus. These steps determine the level of non-uniformity in the image, i.e. the amount of high frequency detail present in the M-filtered image. The level of image non-uniformity can be determined in a variety of ways. Here we stick to a very simple method of evaluating image non-uniformity. First we calculate the Laplacian at every pixel of the M channel output image. Since the DDOG filter is only a linear combination of 4 Gaussians, evaluation of Laplacian from the same implies finally applying a filter which closely resembles Marr's edge detector [Marr, 1982] that also supposedly detects nothing but discontinuities, i.e. non-uniformities in an image. Marr assumed the existence of such filters in the striate cortex also in the form of a

combination of two on and off centered cells arranged in an AND gate structure.

Essentially therefore, in our proposed automated algorithm, there is a first-pass vision at a glance executed by the faster conducting M channel. Detection of non-uniformities in this M channel filtered image (λ) in the striate cortex, leads to a more attentive vision with scrutiny combining the initial M channel picture with the slower conducting P channel filtered image. The extent to which the later picture is combined with the initial picture depends upon the value of λ, i.e. the degree of non-uniformity in the earlier image. The higher the non-uniformity, lesser will be the contribution of the later P filtered picture. On the other hand the more the initial picture is uniform, i.e. devoid of details, the more will be the role of the P channel that consists of the midget cells with higher spatial resolution in the vision with scrutiny.

Such top-down facilitation through an early detection of details/non-uniformity by the initial filtering through M channel also finds justification with respect to higher level object recognition in the works of Moshe Bar and their group [Kveraga et al., 2007] [Bar, 2003]. They used the concept of typical M-biased and P-biased stimuli, low-luminance contrast and achromatic for M and chromatically defined isoluminant for P to demonstrate, by using human neuroimaging techniques, that the stimuli biased towards magnocellular processing were recognized further and

evoked greater activity in the orbitofrontal cortex, associated with object recognition task, compared with the stimuli biased to include primarily parvocellular information. Such M bias was also found to increase the speed between orbitofrontal and inferior temporal cortex in reverse hierarchy. These findings have led them to conclude that fast magnocellular projections to the orbitofrontal cortex, combined with a top-down projection from orbitofrontal to inferior temporal cortex, play a determining role in top-down facilitation of visual object recognition. Based on these works, we may conclude that top-down facilitation possibly follows up the initial M 'scouting', in order to complete the initial gist representation through adding such information that was unavailable in the M output alone.

One may argue that object recognition is a high level vision task, whereas brightness induction is a low level phenomenon. However, even at low level, contemporary anatomical evidences have also been increasingly pointing to possible vertical interactions within the cortical columns in V1 that would provide for a mixture of information originating in parvo and magno LGN layers [Ferrera et al., 1992] [De Valois et al., 2000].

5.5 ACCF model results for various visual stimuli

Below we show some input stimuli and the corresponding brightness profiles of both input and outputs.

5.5.1 White effect

The image in Figure 5.3b shows the output profile of ACCF model when applied to the stimulus in Figure 5.3a, which shows the White effect. The input profile has also been shown for comparison. We can observe that, as expected, the brightness of the patches has shifted in the same direction as the stripes on either side of the patches. This difference in output brightnesses of the patches is more pronounced in ACCF model than in the ODOG model (Figure 5.3c).

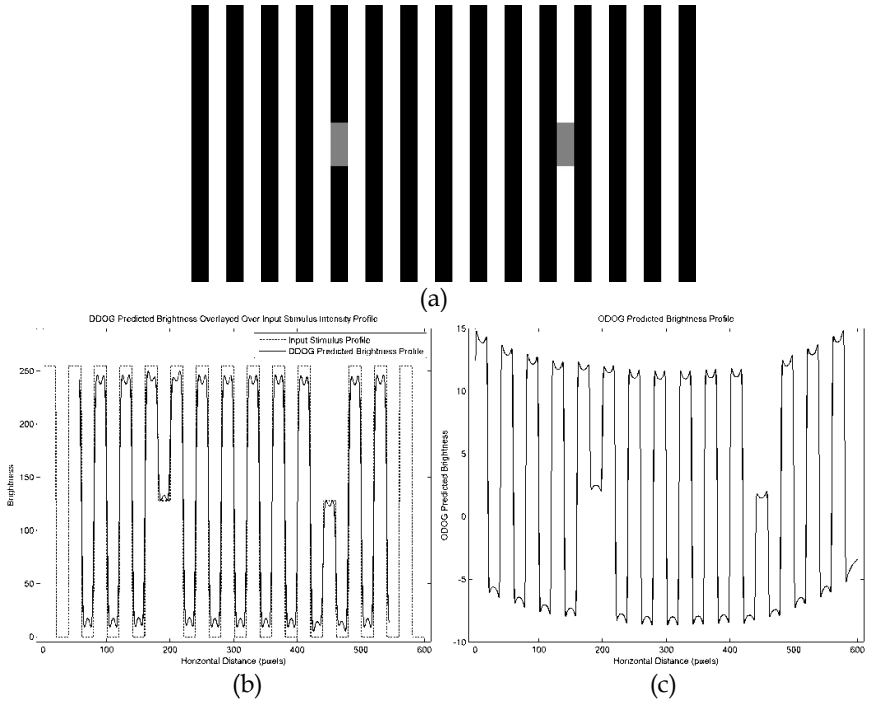


Figure 5.3: (a) Input stimulus for White effect. (b) Output brightness profile (solid) of ACCF model superimposed on the input intensity profile (dotted). (c) Output brightness profile of ODOG model.

5.5.2 SBC

The image in Figure 5.4b shows the output result of ACCF model for the SBC stimulus in Figure 5.4a. As expected we get a brightness-contrast effect, i.e., the patch with a brighter surrounding looks darker while the

patch with a darker surrounding looks brighter. The ODOG output shown in Figure 5.4c also shows the same brightness-contrast effect.

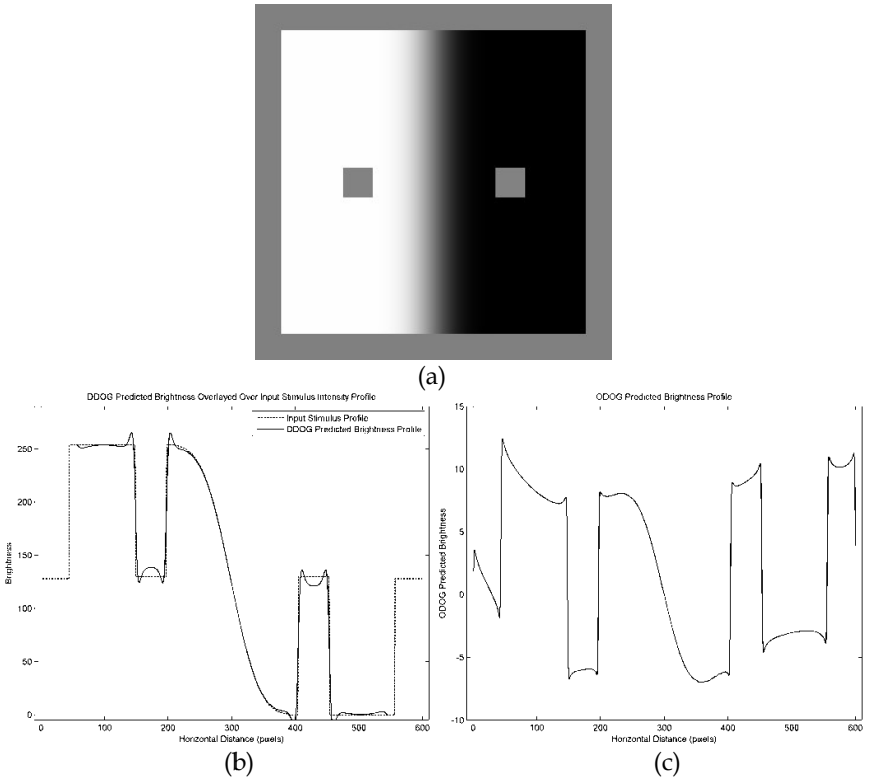


Figure 5.4: (a) Input stimulus of SBC illusion. (b) Output profile (solid) of ACCF model superimposed over the input profile (dotted). (c) Output brightness profile of ODOG model.

5.5.3 Checkerboard

The checkerboard illusion shows a brightness-assimilation effect as seen in Figure 5.5a. The output of ACCF as shown in Figure 5.5b shows an overall shift in patch brightness in the same direction as the surroundings. The ODOG output also shows a shift, but interestingly, not in the right direction. In other words, ODOG here fails to explain the assimilation effect in checkerboard illusion.

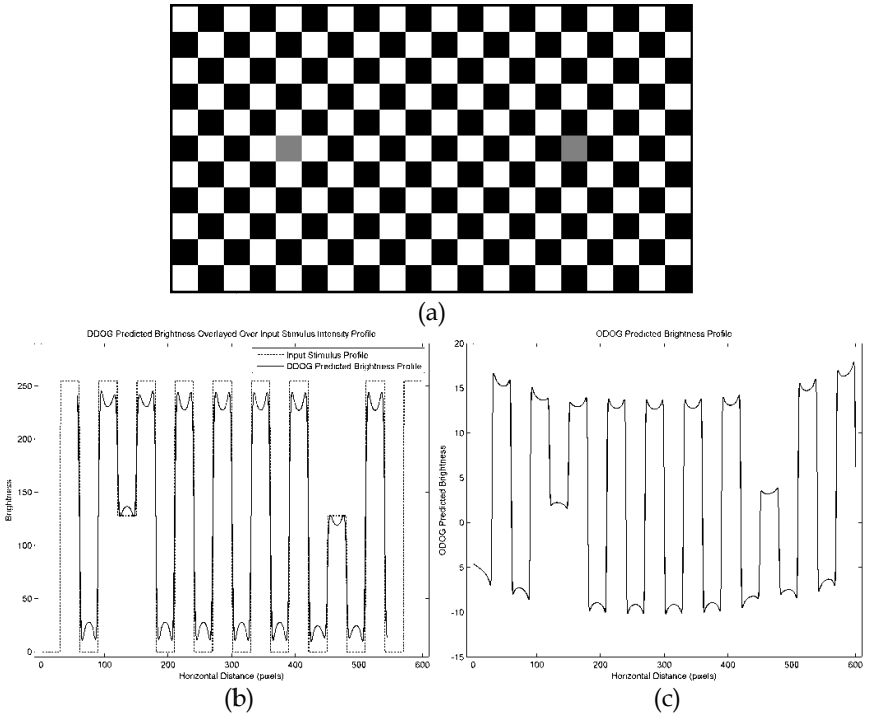


Figure 5.5 (a) Input stimulus of Checkerboard illusion. (b) Output profile (solid) of ACCF model superimposed over the input profile (dotted). (c) Output brightness profile of ODOG model.

5.5.4 Sine grating induction

The Sine grating stimulus as shown in Figure 5.6a consists of a sinusoidally varying background over which a thin grey strip is placed. The grey patch shows an apparent brightness which varies in the opposite direction as the background brightness. The output profile of ACCF model as shown in Figure 5.6b correctly shows the input peaks coinciding with the output troughs and vice versa. The sinusoid being a smoothly varying function has very few edges and as a result the P-filter dominates. The ODOG output shown in Figure 5.6c also shows the same effect.

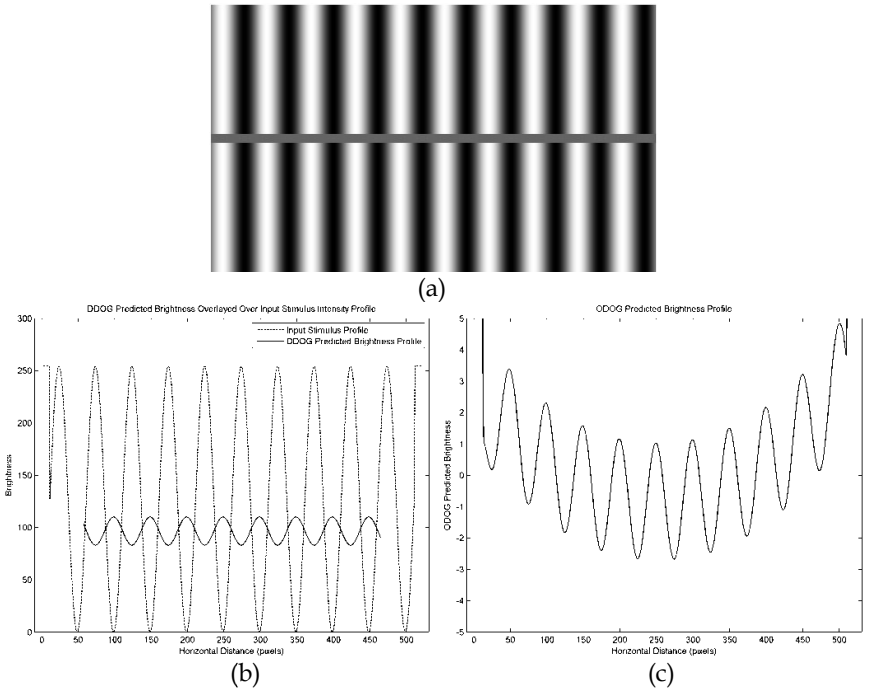


Figure 5.6 (a) Input stimulus of Sine grating illusion. (b) Output profile (solid) of ACCF model superimposed over the input profile (dotted). (c) Output brightness profile of ODOG model.

5.5.5 Square grating induction

The square grating induction is similar to the sine grating except that, instead of a sine wave background it has a square wave background, i.e., the background consists of alternating white and black stripes. Here too the apparent brightness of the grey patch shifts in the opposite direction to the background, but the effect is much weaker than the sine grating illusion. Also unlike the sine grating here the input image is full of edges and therefore the output is mainly formed by the M-filter. Figure 5.7b shows the output profile of ACCF model superimposed on the input profile and as expected, the peaks of the output coincide with the troughs

of the input and vice versa. The ODOG output shown in Figure 5.7c also shows the same effect.

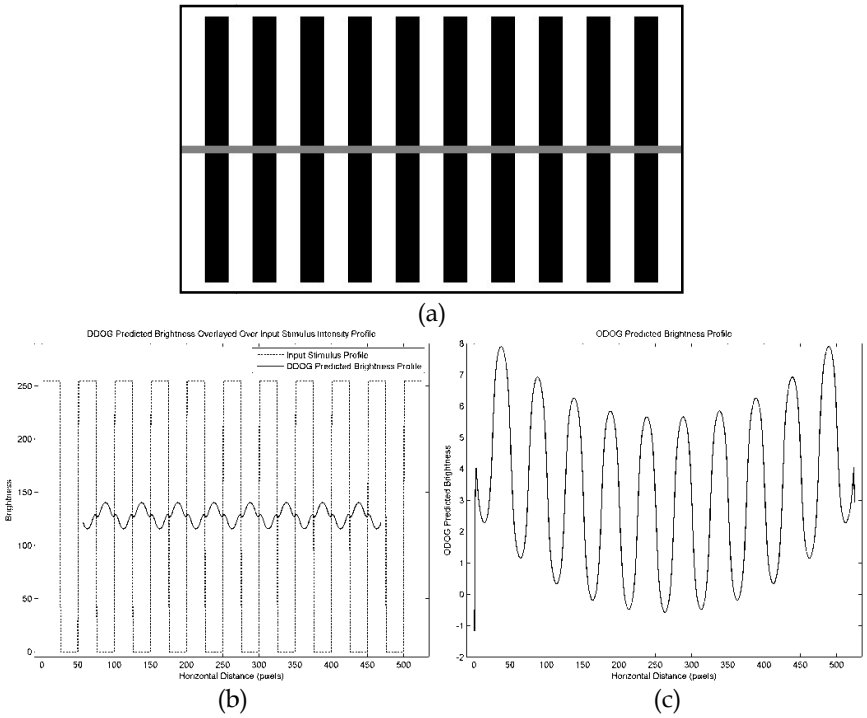


Figure 5.7 (a) Input stimulus of Square grating illusion. (b) Output profile (solid) of ACCF model superimposed over the input profile (dotted). (c) Output brightness profile of ODOG model.

5.5.6 Sine grating to Square grating transition

Figure 5.8a-5.8f show the output profiles of a series of stimuli which show a gradual transition from sine to square grating. The strength of the output response falls gradually from sine to square, as observed. Figure 5.8g shows the relative M-channel weights, $f(\lambda)$, for the sine to square transition. This demonstrates that even though the M-channel weight goes from low to high, the output always shows a brightness-contrast effect, not the brightness-assimilation effect.

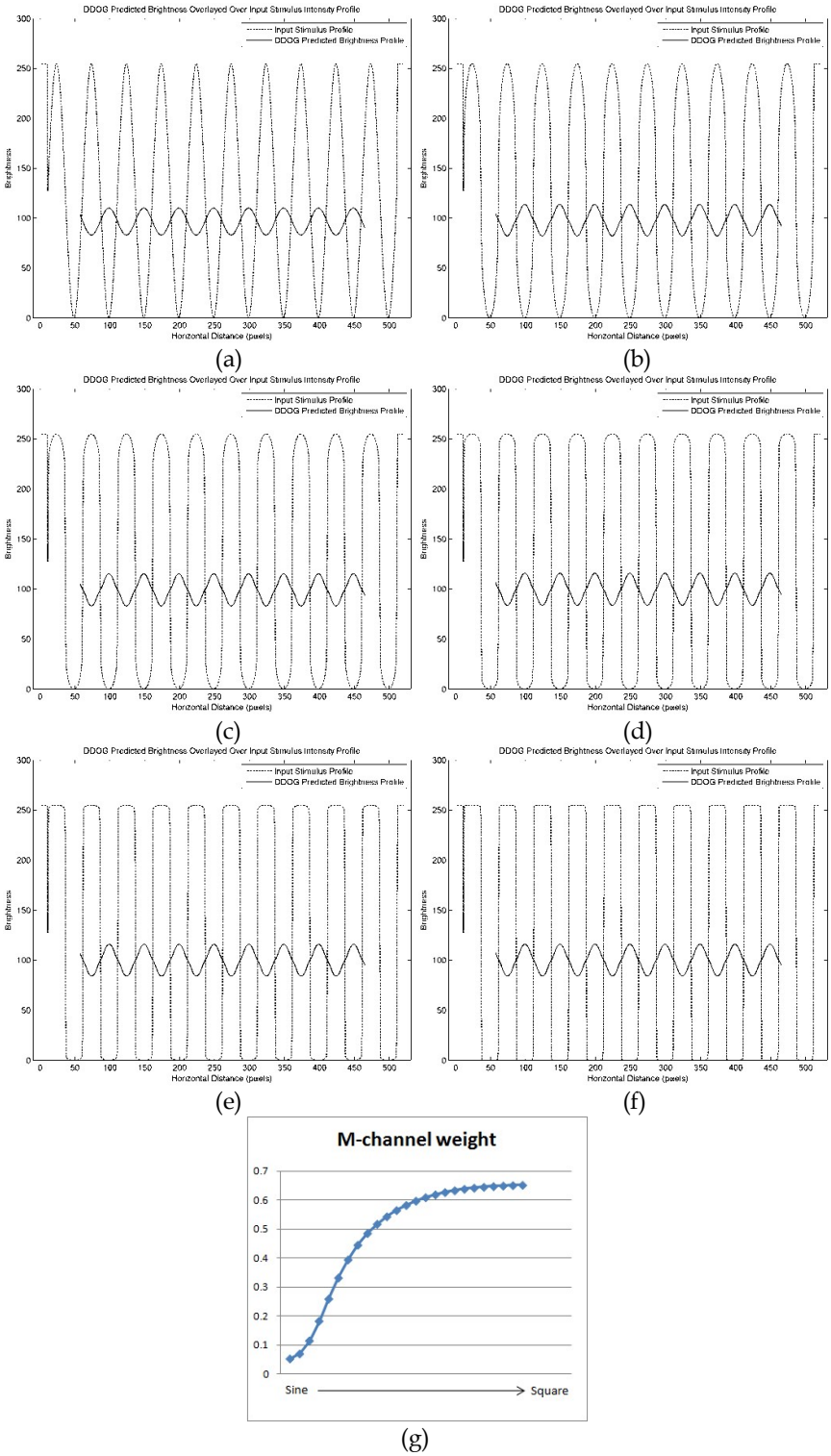
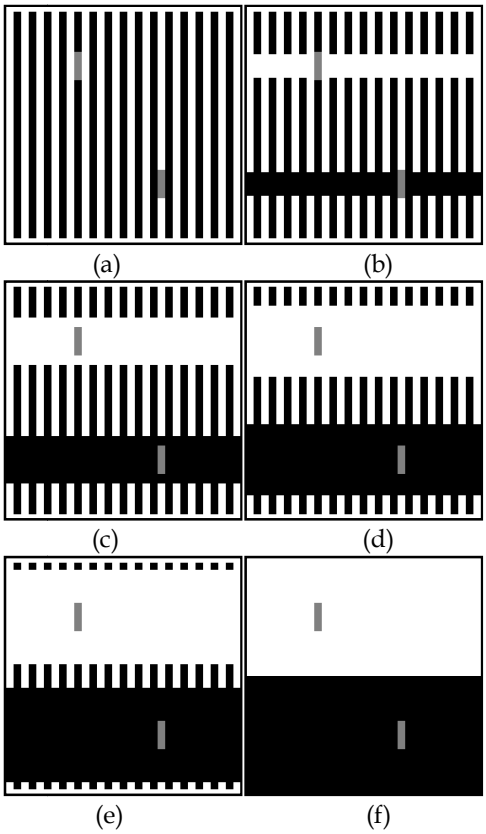
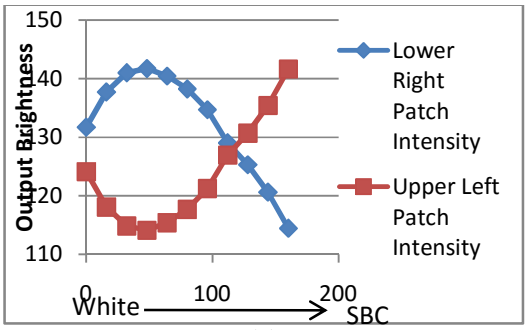


Figure 5.8 (a - f) ACCF output profiles (solid) of Sine to Square transition superimposed over the respective input profiles (dotted). (g) Graph of relative M-channel weights.

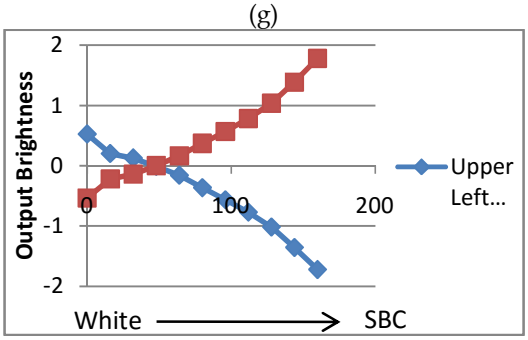
5.5.7 Howe's illusion

Howe's illusion (Figure 5.9b) [Howe, 2001] may be looked upon as an intermediate stage in a gradual transition of the White effect into the SBC illusion. The illusion changes gradually from brightness-assimilation to brightness-contrast. Figure 5.9g shows the output intensities of the two grey patches when the ACCF model is applied to the sequence of images in Figure 5.9a-5.9f. It shows the brightness curve of one of the patches crossing over the brightness curve of the other patch, just as expected from direct observation. The ODOG output (Figure 5.9h) also yields a similar response.





(g)



(h)

Figure 5.9: (a-f) Transition from White effect to SBC, with (b) representing the Howe's illusion. (g) ACCF output intensities of the two grey patches for the sequence of images in (a-f) (h) ODOG output intensities of the two grey patches for the sequence of images in (a-f).

5.5.8 Mach band

The Mach band illusion consists of bright and dark lines seen where a brightness gradient meets a region of constant brightness as shown in Figure 5.10a. The right part of the image is a region of high and constant intensity. The left part of the image is a region of low and constant intensity. The middle part of the image has a brightness gradient which connects the left side to the right side. At the boundary of this gradient region and the right side a bright white line can be seen. Similarly at the other boundary a dark line can be seen. These lines are known as Mach bands. Although the brightness profile of the original image is a non-decreasing function, the apparent brightness profile seems to have

brightness-peaks and brightness-troughs. The output profile of ACCF model shows a brightness-trough and a brightness-peak at the left and right boundaries, respectively, of the gradient region. However the ODOG output does not show such distinct peak or trough. This is another case, apart from the checkerboard illusion, where ACCF model shows a distinct advantage over the ODOG model.

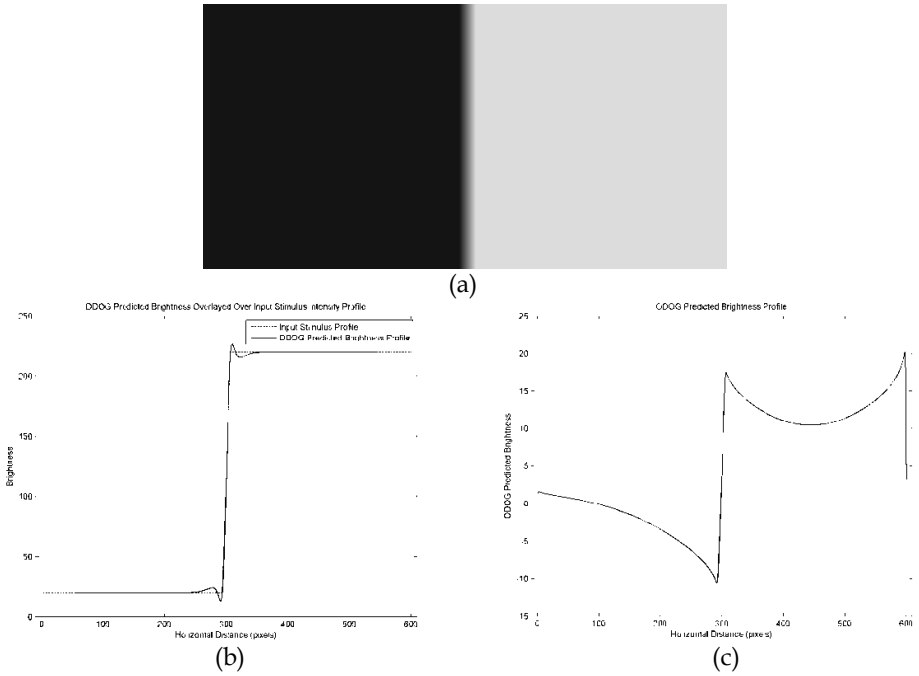


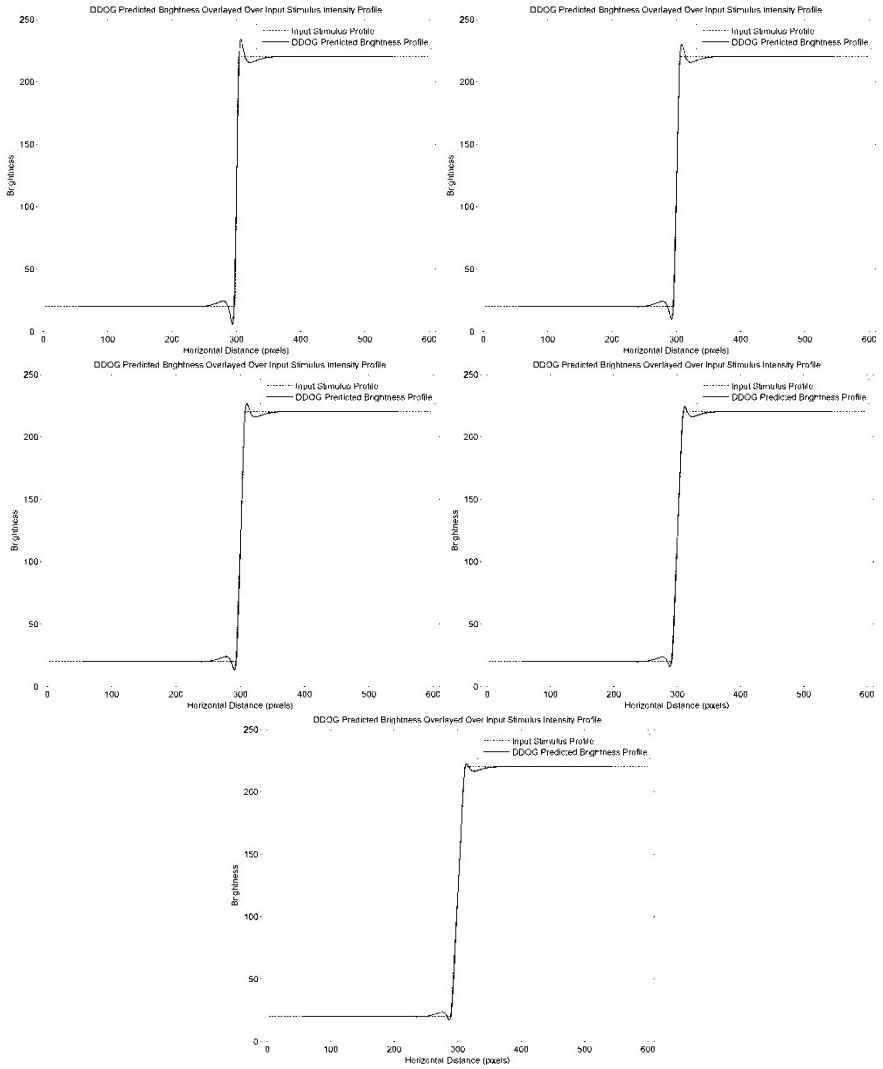
Figure 5.10: (a) Input stimulus of Mach band illusion. (b) Output profile (solid) of ACCF model superimposed over the input profile (dotted). (c) Output brightness profile of ODOG model.

5.5.9 Mach band with varying gradient

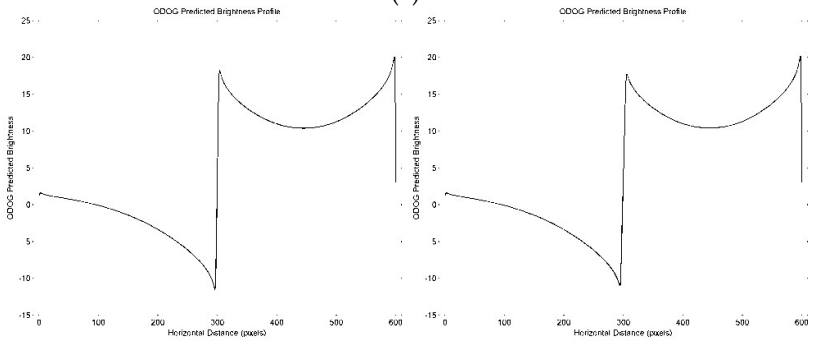
Figure 5.11a shows a series of Mach band images with increasing widths of the gradient region. It is clear and has also been previously shown [Békésy, 1967] [Békésy, 1968] that the brightness of the illusion decreases with increasing width of the gradient region. This fact is clearly reflected

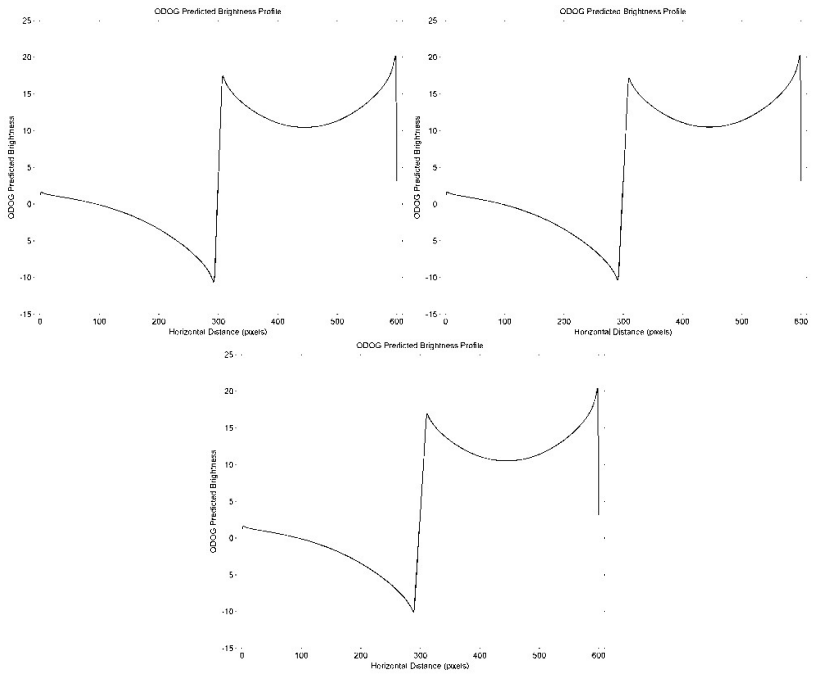
in the output profiles from ACCF (Figure 5.11b) whose peak-size decreases as size of gradient region increases that has been plotted in Figure 5.11d, but the same is not found true for the ODOG output (Figure 5.11c).

(a)

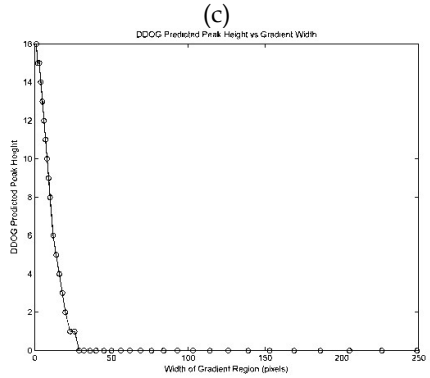

DDOG Predicted Brightness Overlayed Over Input Stimulus Intensity Profile
Input Stimulus Profile
DDOG Predicted Brightness Profile
Brightness
Horizontal Distance (pixels)

(b)


DDOG Predicted Brightness Profile
DDOG Predicted Brightness
Horizontal Distance (pixels)



(c)



(d)

Figure 5.11: (a) Series of Mach Band images with increasing brightness gradients. (b) Brightness profiles for corresponding output images (solid) superimposed over respective input profiles (dotted). (c) ODOG predicted output brightness profile for Mach band stimuli in (a). (d) Graph showing peak height of output profile for the ACCF algorithm plotted against the width of gradient region. As expected it is a decreasing curve.

5.6 Discussion and conclusions

In this chapter we introduced the ACCF model in which we automated the process of estimation of the FOCC (Factor of Contrastive Context), which was left as a free variable in the VCCF model of the previous chapter. We also hypothesized that feedback received by the LGN from the upper

stages of the visual system may be considered as the biological analog of the FOCC, since the LGN is the region where the various visual channels arrive and merge before the visual signals get passed on to the upper stages. We also compared the ACCF algorithm to the well known ODOG model and we demonstrated that while ACCF and ODOG agree in most cases, ACCF does show one benefit compared to ODOG, which is that ACCF is a much simpler model with far fewer number of DOGs than ODOG. While ACCF only uses 4 DOGs, 2 each for the M and P channels, which are isotropic filters, a fact which greatly simplifies its implementation, ODOG uses a set of six oriented DOGs, which are then combined in the last step of ODOG using a nonlinear contrast normalization step. The contrast normalization step weights the output of each of the oriented filters according to their variance and adds the results to produce the final result. It is this contrast normalization step, which is primarily responsible for producing the response to brightness assimilation illusions (e.g. White effect). Although Kingdom [Kingdom, 2011] had pointed out that there is no known physiological analog of the ODOG filter, in this work we have hypothesized how such a nonlinear adaptive combination of different channels may be achieved through a mechanism of feedback from the upper stages of the visual system back into the lower stages. Another feature of the ACCF model is that the DOG-like center-surround response profiles, whose combinations have led to

the DDOG filters, have long been known to be present in neural responses [Rodieck and Stone, 1965]. Furthermore, in the case of the Mach band illusion ACCF gives much better output than the ODOG filter. The ACCF model produces the decreasing nature of the height variation of the Mach band response curve with respect to changing brightness gradients, in accordance with direct perceptual experience. The ACCF model also successfully predicts the checkerboard illusion, unlike the ODOG model. In the ACCF model we can also see the relative weights of the M channel to the P channel. For example, in the sine to square grating transition, it could be seen that the M channel weight increased from very low to high. Even though both sine and square gratings produce a brightness-contrast effect, they are caused by different channels. While the sine grating is closer to the SBC effect, the square grating is much more like the White effect. The ACCF model is also able to reproduce the Howe illusion just like the ODOG model.

Chapter 6

Applications of brightness perception models in image smoothing with edge preservation

In Chapters 4 and 5 we introduced the VCCF and ACCF models which combine the inputs from the M and P visual channels in various possible proportions, in order to model brightness illusions. The M and P channels were individually modeled using the DDOG filter. In this chapter we explore how we can use the DDOG filter in edge preserving noise reduction algorithms.

In this chapter we shall see some practical applications of brightness perception modeling. We will see how brightness perception model can be used to improve traditional smoothing techniques. First we use a lateral inhibition based brightness perception model similar to an adaptive version of the P-channel DDOG filter for the purpose of image noise reduction via image smoothing with edge preservation. In a second application we shall see how a Human Visual System inspired DDOG based M-channel filter performing "Vision at a glance" can produce results comparable to the well known Bilateral Filter, which is a very computationally intensive filter. In this chapter we restrict ourselves only to reducing additive white noise because averaging filters are only good for removing this type of noise by taking advantage of the fact that while

white noise is uncorrelated from one pixel to the next, the image data at neighbouring pixels are highly correlated (except at edge boundaries). The noise reduction algorithms introduced in this chapter are not appropriate for reducing other types of noise, such as salt and pepper noise.

6.1 Introduction

We first introduce various filters used in this chapter.

6.1.1 The Bilateral filter

The bilateral filter is a well known image filtering algorithm primarily used for a wide variety of applications. It was first introduced by Tomasi and Manduchi in 1998 [Tomasi and Manduchi, 1998] as a denoising filter for noise suppression while preserving edges at the same time. Apart from noise removal it is also used for dynamic range compression [Durand and Dorsey, 2002], image cartoonizing [Paris et al., 2009] [Dade] and many others. This makes the bilateral filter a very important algorithm in the field of image processing.

However, despite its importance the bilateral filter is a very computationally intensive algorithm, which keeps it from being used in real time and mobile applications. Several fast alternatives to the bilateral filter have been proposed over the years. Some of these alternatives algorithms are either only perceptual approximations (i.e. produce visually similar results) [Paris and Durand, 2006] or are restricted in the type of spatial and range kernels (defined below) they use. Some

techniques achieve speedup by using polynomial or trigonometric range kernels. Other techniques achieve speedup by using the box function as the spatial kernel.

Bilateral Filter Definition:

The Bilateral filter can be considered as a modification of the Gaussian convolution filter, except that instead of averaging with a fixed set of weights, the weights are modulated according to the intensity values of the pixels that are being averaged. More specifically, the weight corresponding to a neighbouring pixel **n** is increased if its value is close to the value of the centre pixel **c**, and decreased if its value is further away from the centre pixel.

Let I(**p**) be the intensity value of a pixel **p** of the input image I. Here **p** = (i, j) is a two component vector representing the position of a particular pixel located at the position (i , j). I(**p**), a scalar value, is the intensity of the pixel at position **p**. Similarly, let B(**p**) represent the value of the bilateral filtered image at pixel position **p**. Then the Bilateral filtering algorithm can be mathematically expressed as:

$$B(\boldsymbol{c}) = \frac{1}{W(\boldsymbol{c})} \sum_{\boldsymbol{n},\boldsymbol{n}\in N(\boldsymbol{c})} exp\left[-\left(\frac{\|\boldsymbol{n}-\boldsymbol{c}\|^2}{2\sigma_s^2}\right)\right] exp\left[-\left(\frac{|I(\boldsymbol{n})-I(\boldsymbol{c})|^2}{2\sigma_r^2}\right)\right] I(\boldsymbol{n}) \qquad (6.1)$$

Where,

$$W(\boldsymbol{c}) = \sum_{\boldsymbol{n},\boldsymbol{n}\in N(\boldsymbol{c})} exp\left[-\left(\frac{\|\boldsymbol{n}-\boldsymbol{c}\|^2}{2\sigma_s^2}\right)\right] exp\left[-\left(\frac{|I(\boldsymbol{n})-I(\boldsymbol{c})|^2}{2\sigma_r^2}\right)\right] \qquad (6.2)$$

c is the position vector of the center pixel whose bilateral filtered value is being computed. N(**c**) is the set of pixels in the neighbourhood of the center pixel **c** over which the summation is computed and **n** is a particular pixel position in that neighbourhood. σ_s is a parameter called the spatial blur radius and σ_r is called the range parameter.

Equation (1) is nothing but a weighted average computation of the neighbourhood image intensities I(**n**) for each pixel **c** in the image. W(**c**) is a normalization factor defined such that the sum total of the weight factors is unity. This normalization is done so that there is no change in the mean brightness of the image.

In equation (6.1) the first exponential factor inside the summation corresponds to the weight factor in the ordinary Gaussian convolution. This factor is intended to assign lower weights to pixels that are spatially farther away from the central pixel **c**. The second exponential factor is the core feature of bilateral filtering. Its purpose is to assign lower weight values to pixels whose intensities differ greatly from that of the central pixel. The reasoning behind this is that if a pixel value differs vastly from the value of a neighbouring pixel, then they probably belong to different objects in the scene and it makes no sense to average their intensities together. Simple Gaussian averaging of all the neighbouring pixels leads to edges getting blurred in the output image, but if bilateral filtering is used, the strong edges tend to remain sharp. Because of this a bilateral

filter preserves strong edges very well. The selection of filter parameters affects the results significantly and the choice of the range parameter especially has to be done on a case by case basis. Despite producing impressive results the bilateral filter has two major drawbacks. One drawback is that the weak edges in the image also tend to get washed out along with the uniform regions. Another major problem with the bilateral filter is its slow computation speed since, first of all, it cannot be decomposed into two successive 1-D filters, unlike the Gaussian filter. Secondly, the mask being adaptive has to be computed at each pixel. Therefore the convolution operation becomes very tedious. Because of its slowness the bilateral filter is completely impractical for application to video frames in real time, especially when the neighbourhood N($\mathbf{c}$) is of large size.

6.1.2 The DOG filter

The DOG (Difference of Gaussian) function is defined as:

$$\mathrm{DOG}(\boldsymbol{c};\ \mathrm{A}_1, \mathrm{A}_2, \sigma_1, \sigma_2) = \sum_{\boldsymbol{n}, \boldsymbol{n} \in N(\boldsymbol{c})} \left[\mathrm{A}_1 e^{-\left(\frac{\|\boldsymbol{n}-\boldsymbol{c}\|^2}{2\sigma_1^2}\right)} - \mathrm{A}_2 e^{-\left(\frac{\|\boldsymbol{n}-\boldsymbol{c}\|^2}{2\sigma_2^2}\right)} \right] \qquad (6.3)$$

Where, r is the radial distance from the center of the receptive field. The parameters σ_1 and σ_2 signify the widths of the two Gaussian functions.

6.1.3 The DDOG filter

The *Difference of Difference of Gaussians* (DDOG) filter, as introduced in Chapter 4 is defined as:

$$\mathrm{DDOG}(\boldsymbol{c};\ \mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{A}_4, \sigma_1, \sigma_2, \sigma_3, \sigma_4) =$$

$$\mathrm{DOG}(\boldsymbol{c};\ \mathrm{A}_1, \mathrm{A}_2, \sigma_1, \sigma_2) - \mathrm{DOG}(\boldsymbol{c};\ \mathrm{A}_3, \mathrm{A}_4, \sigma_3, \sigma_4) \qquad (6.4)$$

6.2 The Adaptive DOG (ADOG) filter

In this section we use an adaptive version of the ordinary DOG filter to perform noise reduction in color images and then compare it with the Bilateral filter. The DOG filter is very close to the P-channel filter as mentioned in Chapter 4 (Figure 4.3d). So the A-DOG filter may be looked upon as an adaptive version of the P-channel DDOG filter. Biologically the P-channel originates mainly in the foveal region of the eye and is responsible for both color vision and visual acuity. For this reason it makes sense to denoise color images with the A-DOG filter. Our proposed Adaptive DOG (A-DOG) first approximates the gradient at each image point using only the intensity values in a 3×3 region around the image point, just as in the well-known Sobel operator. We calculate the gradient in both directions (x and y). Then the overall gradient magnitude is calculated for each image pixel. Then we use two different DOG filters, one representing a larger receptive field that we apply for gradient values above a threshold and another, representing a smaller receptive field that we apply for gradient values below the threshold, where the threshold is the simple arithmetic mean of the approximated gradient values all across the image. In the real visual system of an organism there may be numerous such filter banks using the billions of neurons in the visual

system representing different pathways. But for simplicity, we have tried to see the effect of a simple two pathway DOG switch which may be looked upon as a first step towards an Adaptive DOG (A-DOG) model of the visual system. The intensity information of the above obtained image is combined with color information from the more blurred DOG image to produce the final image. This step has a physiological justification in terms of the density of rods and cones in the eye. The rods, which can detect only brightness, have an overall higher density in the retina than the cones, which can detect color. So it is justified to take color information from a more blurred image while intensity information has to be taken from less blurred image. Below we represent this algorithm:

6.2.1 Proposed algorithm

- Step 1 - Read the input image.
- Step 2 - Compute the gradient of the image.
- Step 3 – Compute a threshold gradient value by calculating the mean of all gradients.
- Step 4 - Calculate two different DOG filtered images with two different radii. One image is sharper while the other is blurrier. Call the sharper as I_s and the blurry image as I_b.
- Step 5 - For each and every pixel in the image check the following-

  If gradient (pixel) > threshold value

Replace it by the sharper DOG filtered pixel as computed in Step 4 above.

Else

Replace it by the less sharp DOG filtered pixel as computed in Step 4 above.

- Step 6 - Calculate λ for each pixel based on the following formula-

  $\lambda=(R_i+G_i+B_i)/(R_c+G_c+B_c)$

  Where R_i,G_i,B_i are the Red, Green, Blue value of the image resulting from Step 5 (Intensity information) and R_c,G_c,B_c are the Red, Green, Blue value of I_b as calculated in Step 4 (Color Information).

- Step 7 - Calculate red, green, blue value of each pixel for the final output image according to the following formula-

  $R=\lambda*R_c$, $G=\lambda*G_c$, $B=\lambda*B_c$

6.2.2 Results and observations

Our proposed Adaptive Difference of Gaussian (A-DOG) filter provides encouraging results when compared to the bilateral filter [Dasgupta et al., 2013]. It is effective because it preserves the contrast edges even when the spatial smoothing is significantly large and because the time requirement is much lower than bilateral filter, which otherwise provides good results. We demonstrate the results with the help of three images shown below each of which is filtered by four different filtering algorithms, viz. the Gaussian filter (at comparatively lower variance, since higher variances

wipe out the edges), the classical lateral inhibition based DOG filter, the bilateral filter and finally our A-DOG filter.

(a)

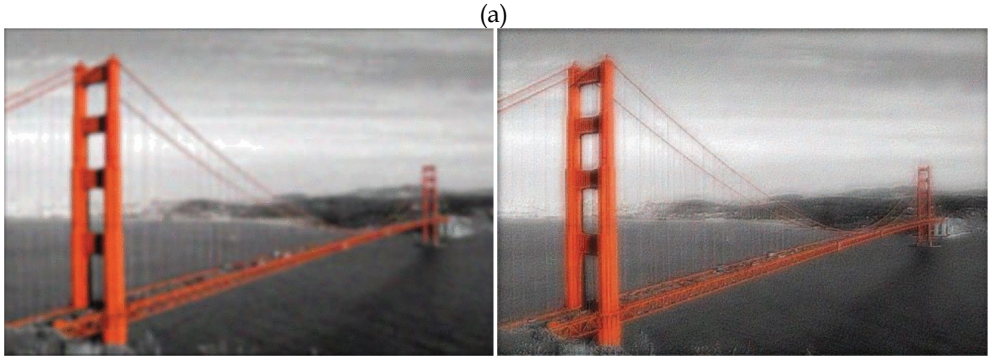
(b)

(c)

(d) (e)

Figure 6.1: (a) Original noisy image (Image 1) (b) Simple Gaussian filtered Image (c) DOG filtered Output (d) Bilateral Filtered Image (e) Proposed A-DOG Filtered Image

(a)

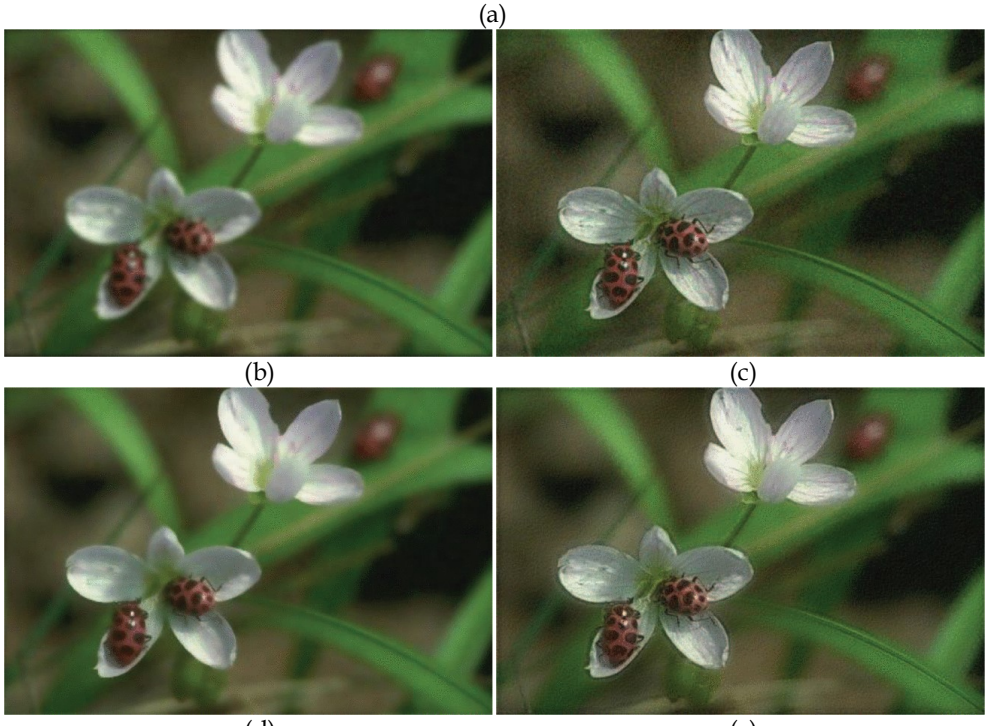


(b) (c)

(d) (e)

Figure 6.2: (a) Original noisy image (Image 2) (b) Simple Gaussian filtered Image (c) DOG filtered Output (d) Bilateral Filtered Image (e) Proposed A-DOG Filtered Image

(a)



(b) (c)

(d) (e)

Figure 6.3: (a) Original noisy image (Image 2) (b) Simple Gaussian filtered Image (c) DOG filtered Output (d) Bilateral Filtered Image (e) Proposed A-DOG Filtered Image

A visual comparison of four different filtering techniques for three different images is shown in Figures 6.1a to 6.3e. Although the bilateral filter gives impressive results, some creases, such as on the leaf in Figure 6.1a, are better preserved in our proposed technique.

It can be clearly seen that like the bilateral filter, the A-DOG filter is better than a simple Gaussian filter, because the Gaussian filter is unable to preserve significant details of an image. Similarly our algorithm also has an advantage over the simple DOG filter because the DOG operator by itself is not capable of deciding which edges to preserve and which edges to omit because it lacks a mechanism to prioritize edges. It only globally enhances the edges and hence does not perform well for noisy images. In our algorithm we used the gradient value which is the indication of an edge and from the gradient we can easily distinguish the stronger edge from a weaker edge. So, for strong edges we suggest to use a DOG filter with larger radii which will less disturb the strong edges and for weak edges the filter with smaller radii which will preserve some of these details. So, automatically we are keeping strong edges as well as some weak edges intact. Moreover, we are taking color information from the less sharp DOG filtered image and intensity information from the A-DOG filtered image. Now, the threshold value will determine how much details we want to preserve. So, by changing the threshold value we can actually control the amount of details we need in the filtered image.

6.3 Using the DDOG based M-channel filter for image smoothing with edge preservation

As explained in Chapter 4 the M-channel filter can be considered as performing "vision at a glance" in the human visual system (HVS). "Vision at a glance" may be defined as the high-level generalized, categorical scene interpretation, identifying "forest before trees". The established and prevalent viewpoint conceives that the human visual system perceives categorical information at a glance (using high level cortical mechanisms) and misses (or assumes) details that are detected by lower areas but not represented in the individual high-level receptive fields [Ghosh, 2012]. We perceive the details later by focusing serially on the components and features, slowly scanning them one at a time [Bullier, 2001]. The Reverse Hierarchy Theory (RHT) [Hochstein and Ahissar, 2002] proposes that such an initial "vision at a glance" includes results of automatic and implicit bottom-up processing, which makes the initial explicit perception introspectively direct without conscious antecedents.

In Chapter 4 we used the DDOG filter (equation 6.4) for the purpose of modeling the M & P channels of the human visual system. In the previous section we used the Adaptive DOG filter to show that it could achieve similar levels of edge preservation quality as the bilateral filter. In this section, instead of using the DOG we use the M channel filter of chapter 4 for performing image smoothing with edge preservation. The

M-channel filter has a image smoothing property because it carries less detail about the visual scene and therefore is spatially smoother than the P channel, while also being temporally faster, which is why it is the channel used as a first estimate for "vision at a glance". But because of the extra Gaussians of the DDOG, it also has an edge preservation property. In this section we compare the performance of the M-channel filter with the well-known bilateral filter in achieving such "vision at a glance" which is akin to image preprocessing in the computer vision domain. We find that at higher noise levels, the M-channel filter performs better than the bilateral filter in terms of reducing noise while preserving edge details. The M-channel filter is also significantly simpler and therefore faster than the bilateral filter. Overall, the M-channel filter enables us to model, simulate and arrive at a better understanding of some of the initial mechanisms in visual pathway, while simultaneously providing a fast, biologically inspired algorithm for digital image preprocessing.

The M-channel filter possesses the characteristic of suppressing the high frequency information and provides a holistic view containing less detail (but enough for the initial sensorial percept), to achieve the goal of providing "vision at a glance". In chapter 4 the coefficients of the DDOG equation 4 for the case of the M-channel filter were as shown in Table 6.1:

	A_1	A_2	A_3	A_4	Sampling Interval
M-filter	10	0.5	0.5	0.08	0.5

Table 6.1

With the values of $\sigma_1=0.7$, $\sigma_2=3\sigma_1$, $\sigma_3=3\sigma_1$, $\sigma_4=9.3\sigma_1$ for the filter sizes.

6.3.1 Comparison of M-channel filter and Bilateral filter by visual inspection

In this section we compare the performance of the DDOG based M-channel filter with the bilateral filter on a grayscale image by visual inspection. We present the results of filtering on a standard benchmark image Lena (Figure 6.4a) with zero-mean, Gaussian white noise of variance 0.01 added to it (Figure 6.4b). The results of filtering are shown Figure 6.4c and 6.4d. Figure 6.4c and 6.4d show the effectiveness of the M-channel and Bilateral filters respectively as filters helping in 'vision at a glance'. It can be observed that both of them filter out the noise and provide a smoothing effect when compared to Figure 6.4b. We can also observe that the edge details have been preserved. For example, looking at the lining of the hat in the region marked by the white rectangle, we can observe that it has been preserved in the filtered images when compared to the original image Figure 6.4a. Therefore, these qualitative results, as well as some quantitative ones [Bhattacharjee et al., 2015] motivate us in future to conduct detailed experimental studies and subsequent analysis

of the performance of M-channel vs. bilateral filter for achieving 'vision at a glance'. Our studies conducted so far reveal several advantages of using the DDOG filter over bilateral filter which has been detailed in the Appendix of the present thesis after the concluding chapter.

Figure 6.4 (a) Lena.jpg 512 x 512 grayscale image. (b) Lena image with additive white Gaussian noise of variance 0.01. (c) M-channel filter result on Figure 6.4(b). (d) Bilateral filter (spatial-domain standard deviation = 3 and the intensity-domain standard deviation = 0.1) result on Figure 6.4(b). The region of interest has been marked using a white rectangle in each figure.

6.4 Conclusion

In this chapter we explored the effectiveness of the physiologically inspired DDOG filter, which we hypothesized to model the M and P channels of the visual system, for the purpose of designing image smoothing algorithms that preserve edges. We compared our algorithms with the bilateral filter which is well known for preserving strong edges while smoothing out relatively uniform regions.

The Adaptive Difference of Gaussian (ADOG) algorithm is simpler and faster compared to bilateral filtering, but gives output of comparable quality. In contrast to the other DDOG based algorithm introduced in this chapter the ADOG algorithm uses the simpler DOG filter instead of the DDOG filter. This is because the DDOG filter of the P channel introduced in Chapter 4 is a very close approximation of the DOG filter. The P channel is responsible for high acuity color vision. Therefore we test this algorithm on color images.

The ADOG filter utilizes the fact that although the human eye is more sensitive to color noise than intensity noise it takes spatial information mainly from the intensity signal. Therefore the ADOG filter takes color information from more blurred image and intensity information from the less blurred image of the same object. Moreover it is adaptive in nature as it is dependent on gradient of the edges. So our algorithm is capable of enhancing edges by taking into account gradient, intensity and color information. This algorithm is also capable of

smoothing the image. So smoothing as well as edge enhancement can be achieved simultaneously by our algorithm. Being faster than the bilateral filter, it may be suitable for real life applications to image enhancement.

In this chapter, we have also used the Human Visual System inspired DDOG based M-channel filter for modeling 'vision at a glance', that closely resembles image denoising with edge preservation, an important task in computer vision. We compared our algorithm with the bilateral filter and from the results we find that the performance of both the filters as satisfactory in achieving image smoothing. The ability to pre-compute the filtering kernel in case of M-channel filter helps in very fast implementation of the filter. On the other hand, bilateral filtering can smooth images to differing extents, while preserving the edge details by appropriate setting of the range parameter of the spatial and range kernel. The highlight of M-channel filter is the simplicity of the kernel and its low computation requirements which make it suitable for initial processing of an image. These factors in favour of the DDOG filter over bilateral filter have been detailed in the Appendix of this thesis.

Chapter 7

In lieu of a conclusion

In this concluding chapter we first briefly outline the findings and contributions of the present thesis to brightness perception through the adoption of a spatial filtering approach, and then we discuss some shortcomings of spatial filtering models of brightness perception in general. We show some experimental observations which raise fundamental questions about the validity of spatial filtering algorithms. Finally, we speculate on what future directions could be taken to resolve these difficulties.

We begin this chapter by first summarizing the basic contributions of the present thesis. In this thesis we investigate brightness perception from the perspective of spatial filtering. The thesis comprises of both experimental and theoretical contributions. At first, we have experimentally demonstrated the limitations of existing spatial filtering models, especially the well established ODOG model. We then proposed the *Varying Contrastive Context Filter* (VCCF) and the *Adaptive Contrastive Context Filter* (ACCF) models respectively, based on a new spatial filter namely the *Difference of Difference of Gaussians* (DDOG) filter, as alternatives in order to address some shortcomings of the ODOG model. Next, some applications of the alternative models to image noise reduction with edge preservation which is a fundamental challenge in computer vision, were shown.

In the remaining part of this concluding chapter, we would traverse beyond the domain of spatial filtering by discussing some experimental observations on the scaling behaviour of various brightness stimuli, viz. the Mach band stimulus, the Sine and Square grating stimuli and the Hermann grid stimulus. These experiments will demonstrate the various complexities of brightness perception modeling and the limitations of spatial filtering models of brightness perception.

After that we demonstrate an interesting quirk in the perception of the Hermann grid illusion which is difficult to explain using spatial filtering algorithms.

7.1 Scaling behavior of visual illusions

By the term scaling behavior we mean the changes in the illusory effect brought about by a change in one of the length-scale parameters of the illusion. For example, in the Shifted-White stimulus, one scale parameter would be the width of the alternating black and white bars or equivalently its inverse, i.e. the number of bars per unit length, also known as spatial frequency. It is well known that the perceived brightnesses of the grey patches depend on the width of the bars. A second scale parameter would be the vertical size of the grey patch. We have already seen in Chapter 3 that this second scale parameter leads us to some limitations of the ODOG model.

7.2 Scaling behavior of the Mach band illusion

The Mach band illusion, as introduced in Section 1.3.9, shows a very interesting scaling behavior, which was partly discussed in section 3.3, that

defies the predictions of many brightness models. The Mach band stimulus only possesses a single length scale parameter, which is the width of the gradient region. Peculiarly, the illusory effect is not visible all over the gradient region, but only at the edges where there is a change of the gradient. However, there is no scale parameter attached with the gradient change as it occurs instantly at a single position on the x-axis. Interestingly, as shown in Figure 7.1, the Mach band illusion scales proportionally with the length parameter.

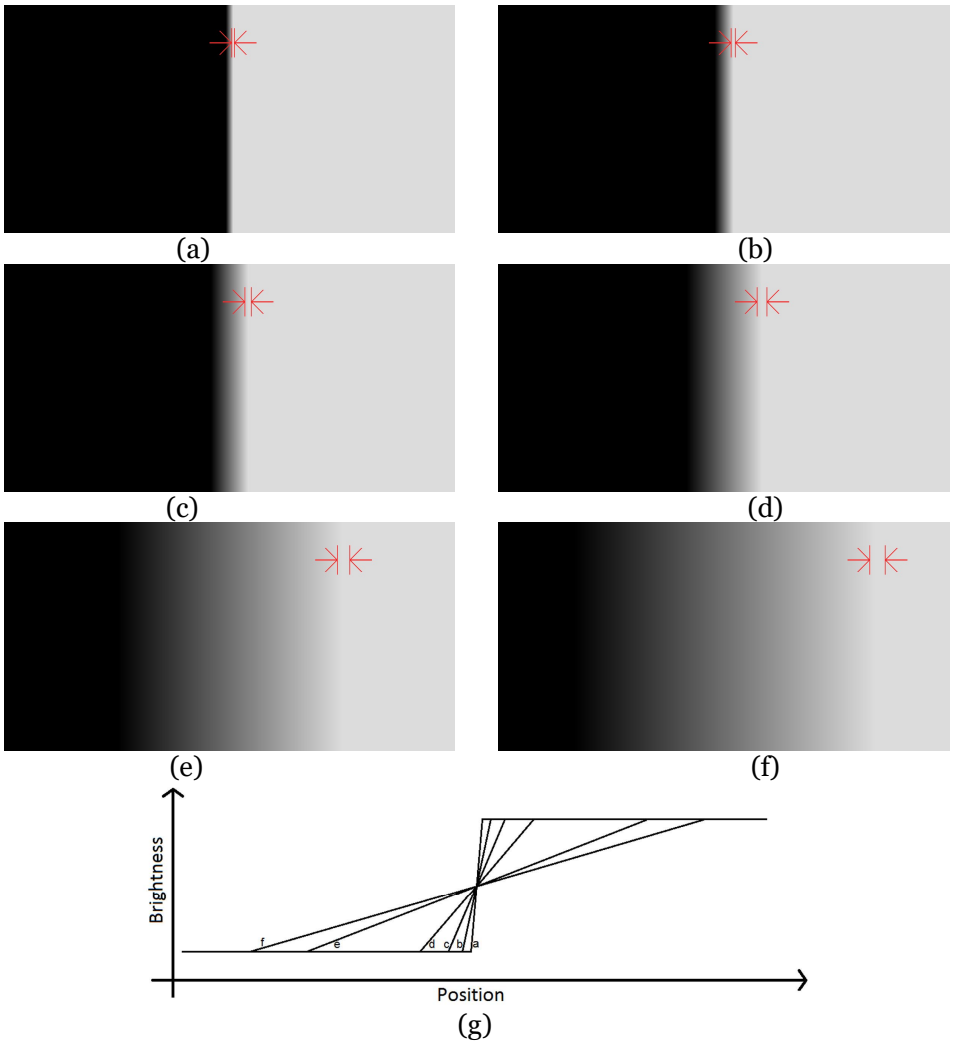


Figure 7.1: Mach band stimuli for various widths of the gradient region. Figures (a) through (f) have their regions of gradient progressively increased in size in ratios of 2:5:10:20:60:80 respectively, i.e. the horizontal gradients have been scaled down in that proportion as shown in (g).

To experimentally test this claim we conducted an experiment to measure the width of the Mach band as a function of the size of the gradient region. In this experiment a set of volunteers were shown a sequence of images with increasing size of the gradient region as in the demonstrations shown in Figure 7.1a to 7.1f. The volunteers were instructed to enclose the white Mach band between a pair of cursors by moving the cursors left or right in a GUI software as shown in Figure 7.2. The cursors must be brought as close as possible while still enclosing the Mach band completely. The distance between the cursors is then taken to be the width of the Mach band. When the average distance between the cursors is plotted against the size of the gradient region (both being measured in the same unit), we get a curve as shown in Figure 7.3.

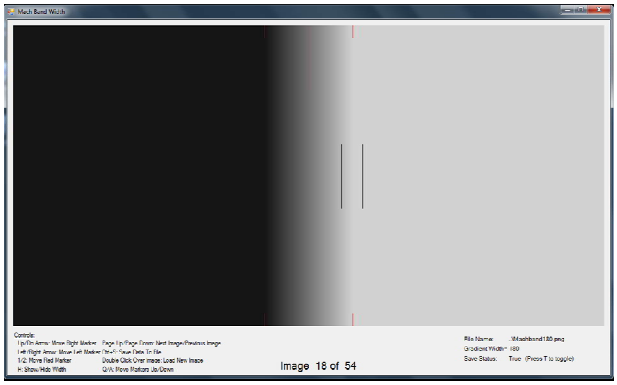


Figure 7.2: Screenshot of the GUI presented to volunteers for the experimental measurement of the width of the Mach band as a function of the size of the gradient region.

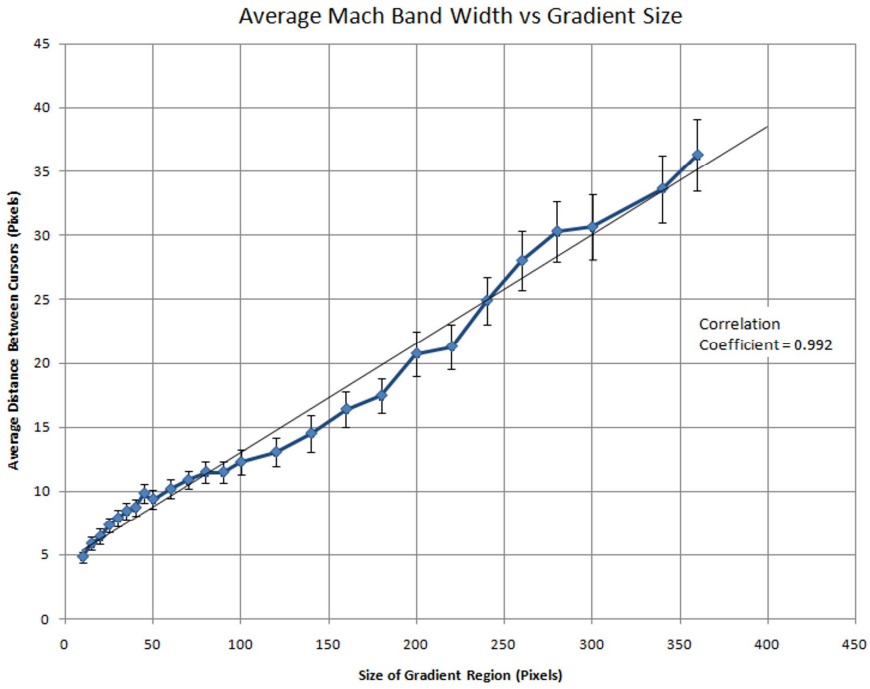


Figure 7.3: Mach band width measured as a function of the size of gradient region averaged over nine different volunteers. Error bars show the standard errors of the means.

The next image demonstrates that the width of a Mach band is a function of the gradients only and does not depend on the absolute levels of brightness at a particular point.

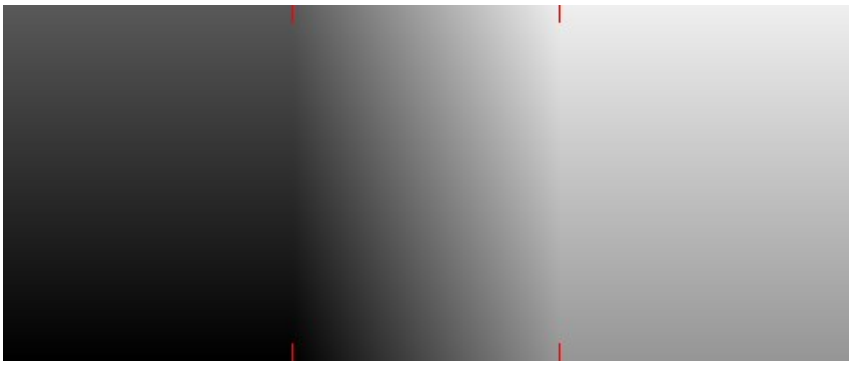


Figure 7.4: Every horizontal line in this image has the same brightness profile as that of the horizontal line at the bottom of the image, except for an added offset. The offset itself increases as we move from bottom to top.

Figure 7.4 has been generated such that the gradient of the horizontal brightness profile is independent of the vertical position. Inspection of Figure

7.4 will reveal that the width of the Mach band is constant throughout from top to bottom. Therefore it can be concluded that the width of the Mach band is a function only of the intensity gradients on the either side of the band and not dependent on the absolute brightness levels themselves. Therefore

$$width = f(S_L, S_R) \qquad (7.1)$$

Where S_L is the gradient on the left side and S_R is the gradient on the right side of the band. If the image is inverted left to right then the width remains unchanged. Therefore

$$f(S_L, S_R) = f(-S_L, -S_R) \qquad (7.2)$$

The widths of the clearly visible Mach bands seem to be scaled up proportionally in size. Mathematically this can be written as

$$f(kS_L, kS_R) = \frac{1}{k} f(S_L, S_R) \qquad (7.3)$$

At smaller scales the bands are thin and sharp while at larger scales the bands are wide and less prominent. Note particularly that at the smallest scale the Mach band is hardly visible since the band itself has been compressed into a very thin line. This is expressed by

$$\lim_{|S_L| \to \infty} f(S_L, S_R) = 0 \qquad (7.4)$$

This explains why no Mach band can be seen for a step change of intensity.

In the context of our previous assumption that the width of a Mach band is a function of the neighbourhood gradients only, this scaling property implies that

$$f(S_L, 0) \propto \frac{1}{|S_L|} \qquad (7.5)$$

Also if the gradients on both sides are nearly equal then the width is very large. Therefore as $S_L \to S_R$, $f(S_L, S_R) \to \infty$.

$\lim_{S_L \to S_R} f(S_L, S_R) = \infty$ (7.6)

The above two equations can be combined in the single equation

$width = f(S_L, S_R) \propto |S_L^n - S_R^n|^{-1/n}$ (7.7)

Where n is a real positive number. Although this last relation cannot be strictly derived from the previous relations but such a power law seems to be the simplest function which satisfies all of the above properties while containing a minimum number of coefficients.

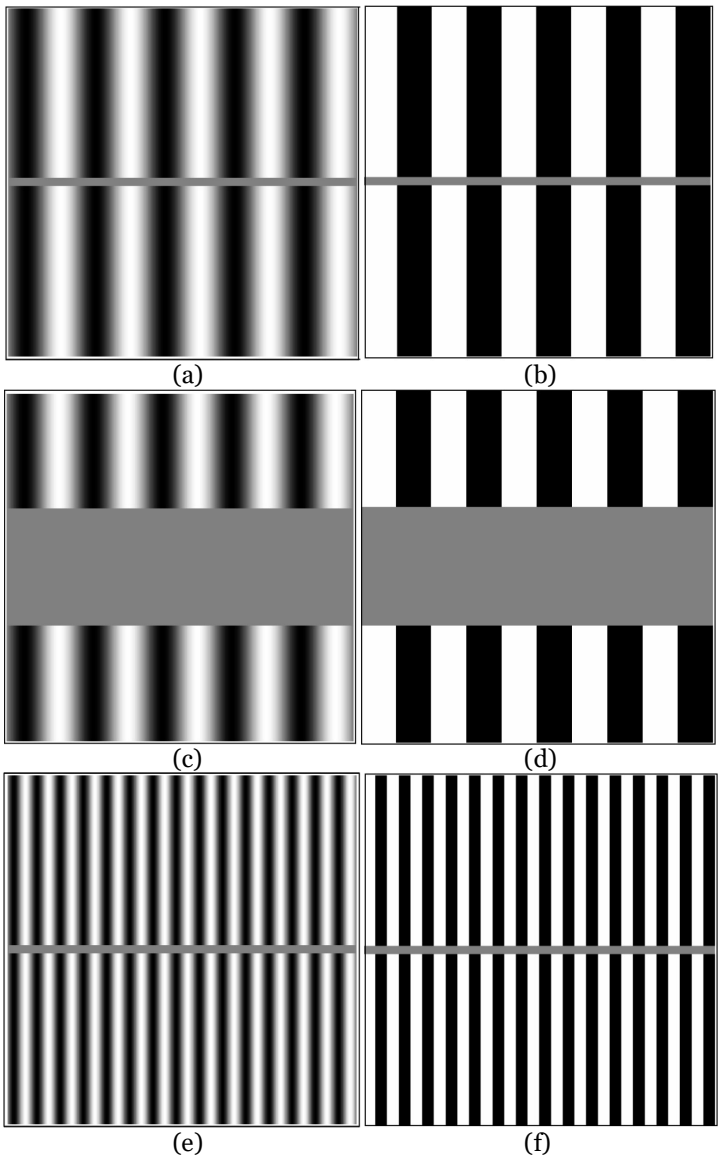
The above results reveal some problems with spatial filtering algorithms that may be used to explain brightness illusions. These problems may be stated in brief as follows:

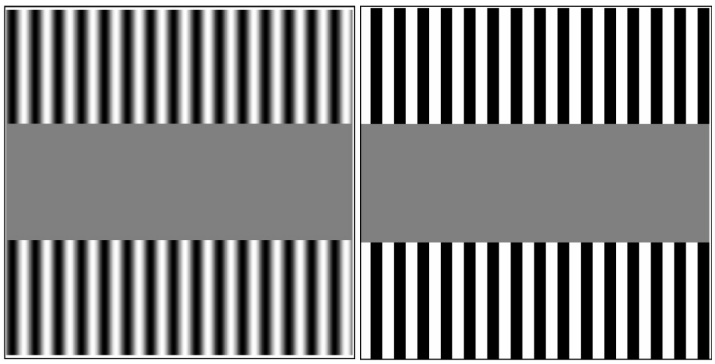
Any physically realizable filter function must have a finite, non-zero width with some typical length scales, e.g. the length scales of the DOG filter are characterized by the parameters σ1 and σ2 of Equation 2.2 of the DOG filter function. The response signals would therefore also have these same length scales. Consequently, very short (almost tending to zero) or very long responses as we obtain here, cannot be obtained by linear filtering.

7.3 Scaling behavior of the Sine grating and Square grating illusions

In this section we describe certain experimental observations to reveal the scaling behavior of sine and square grating illusions [Bakshi et al., 2017]. This experiment was conducted upon volunteer subjects. Sine and square grating stimuli, such as shown in Figures 7.5a to 7.5h, were shown to the subjects at various cycle frequencies. Cycle frequencies ranged from 20

cycles/radian to 121 cycles/radian in case of sine-wave grating. For the square-wave grating, minimum frequency was 20 cycles/radian and maximum frequency was 101 cycles/radian. The size of the entire illusion was 0.15 radian by 0.15 radian and it was same for both the gratings. Grayscale intensity value of the test strip was 150. Before the experiments, the subjects were instructed to concentrate on the approximate horizontal centerline of the strip and report the perceptibility of the illusion at the centerline, i.e. whether or not brightness undulations could be seen along the horizontal centerline of the gray-strip.


(a)
(b)
(c)
(d)
(e)
(f)



(g) (h)

Figure 7.5: (a) Low frequency sine-wave grating stimulus with narrow gray strip. Illusory brightness undulations are visible all throughout the gray-strip. (b) Low frequency square-wave grating stimulus with narrow gray strip. Illusory brightness undulations are visible all throughout the gray-strip. (c) Low frequency sine-wave grating stimulus with wide gray- strip. Illusory brightness undulations are visible only along the upper and lower edges of the gray-strip. (d) Low frequency square-wave grating stimulus with wide gray-strip. Illusory brightness undulations are visible only along the upper and lower edges of the gray-strip. (e) High frequency sine-wave grating stimulus with narrow gray-strip. Illusory brightness undulations are visible all throughout the gray-strip. (f) High frequency square-wave grating stimulus with narrow gray- strip. Illusory brightness undulations are visible all throughout the gray-strip. (g) High frequency sine-wave grating stimulus with wide gray strip. Illusory brightness undulations are visible only along the upper and lower edges of the gray-strip. (h) High frequency square-wave grating stimulus with wide gray strip. Illusory brightness undulations are visible only along the upper and lower edges of the gray-strip.

For any specific value of cycle frequency of grating, we started with a width of .002 radian of the gray-strip. Then we increased the width in steps until the gray strip fully overlaid the grating and marked that strip width at which response of the subject reverses (response changes from 'yes' to 'no' or 'no' to 'yes'). From this value of strip width, at which the transition in response occurred, the absolute threshold for strip width was determined for a particular frequency. This procedure was done for each spatial frequency of both sine-wave and square-wave grating. During the experiment the width of the strip was being changed using key presses. For each key press the width of the strip was being changed by 0.0005 radians. To avoid tiring the subject, the

interval between two successive key presses was kept to more than 10 seconds.

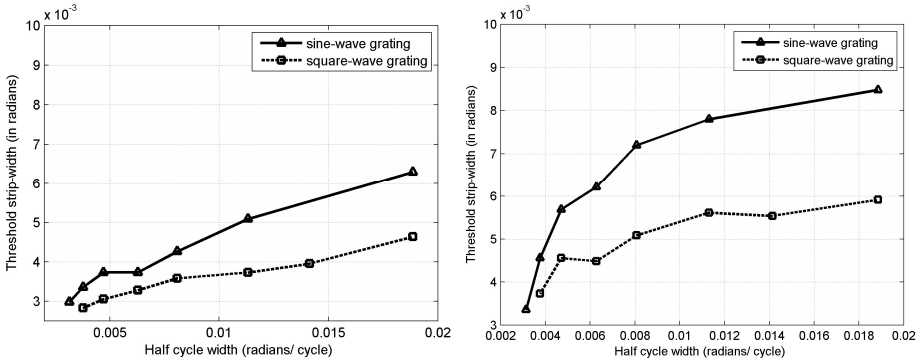


Figure 7.6: (a) Threshold strip-width vs. spatial frequency curve for subject 1. (b) Threshold strip-width vs. spatial frequency curve for subject 2.

The results from the experiment 1 have been demonstrated in Figure 7.6a & 7.6b. From these figures it can be distinctly seen that for every subject and at each spatial frequency of the grating there exists a threshold width of the gray test strip above which the illusory brightness undulations is not visible at the centerline. This implies that the illusory effects in the Grating Induction illusion are confined to the boundaries of the gray-strip. Even though the values of thresholds at different frequencies can vary from subject to subject, some threshold must exist for every subject.

These results show that as the spatial cycle-width of the grating increases the threshold also increases. From Figures 7.6a & 7.6b, it can be seen that the visibility threshold of the illusory effect increases roughly linearly with the wavelength of the background sine/square wave. This implies that the length scale of the illusory effect increases in proportion with the wavelength of the background waveform. This suggests that the perceptual mechanism which is generating this illusion is capable of generating the

illusory response at all length scales from very small to very large, which as mentioned in the previous section about the scaling properties of Mach bands, is difficult to explain using spatial filtering models, since spatial filters can only produce responses of limited length scales both on very large length scales and very short length scales.

We can also see that irrespective of the subject, at any given wavelength of the background waveform the threshold value for square-wave grating is smaller than that for sine-wave grating. This clearly indicates, in agreement with the previous assertions [McCourt and Foley, 1985], that the sine-wave grating induction illusion is much stronger than square-wave grating induction illusion.

7.4 Scaling properties of the Hermann grid illusion

In this section we describe an experiment to measure the scaling properties of the Hermann grid illusion using stimuli as shown in Figure 7.7a-7.7c [Bakshi et al., 2017]. The subjects were shown several sequences of Hermann grid illusions which differed from one another by the size of their squares. Maximum side-length of squares was taken 0.03 radian and the minimum count was 0.008 radian. In this range we varied the side-length of the squares in the grid.

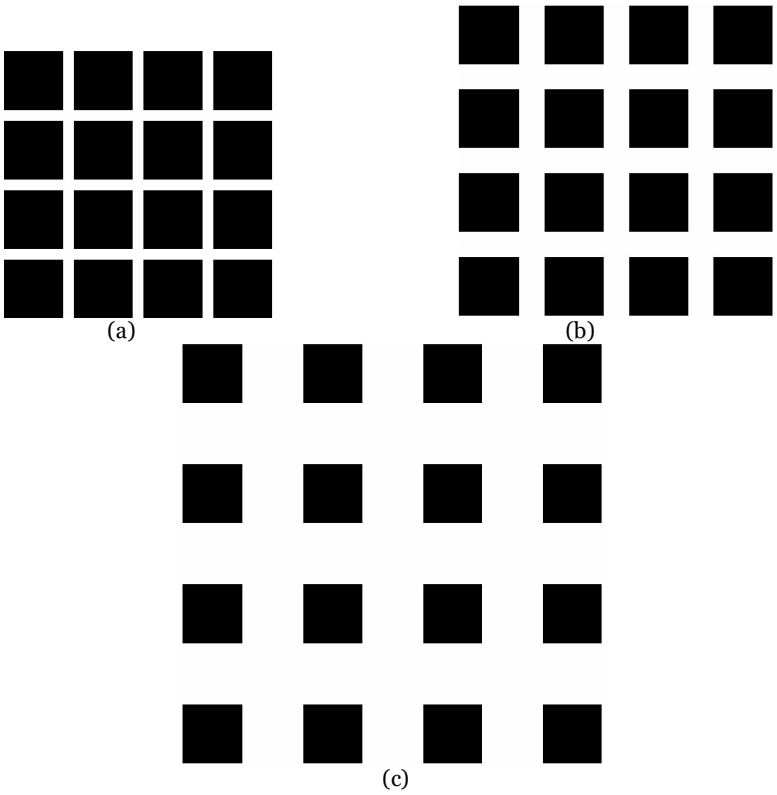
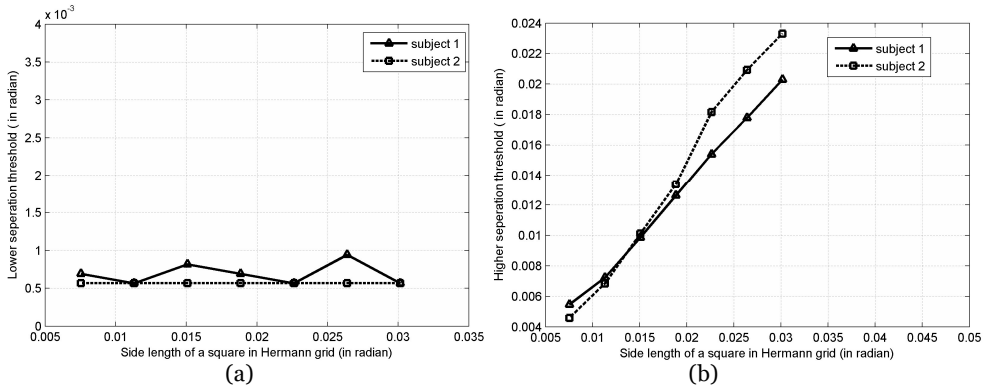


Figure 7.7: (a) The Hermann grid stimulus with small separation width. Illusory blobs are prominently visible at the intersections. (b) The Hermann grid stimulus with medium separation width. Illusory blobs can be seen at the intersections, but are less prominent than Figure (a). (c) The Hermann grid stimulus with large separation width. No illusory blobs can be seen anymore.

We started with a Hermann grid with specific square side-length and asked the subject whether he could see the illusory blobs or not. Initially, the width of separation between the squares was taken to be 5×10^{-4} radians. Then we increased the separation width between the squares in small steps, keeping the square-size constant, as shown in Figures 7.7a-7.7c, while recording the perceptibility of the illusion by the subject. At each step, if any transition in subject response occurred we noted the separation width and used it for the separation width threshold calculation. We repeated this procedure using Hermann grids with squares of different side-lengths. Separation width was

incremented in steps of $5x10^{-4}$ radians. The time interval between two steps was maintained above 15 seconds in order to avoid subject fatigue.

It was observed, that below a minimum separation between squares, the illusory effect of Hermann grid was not perceivable. We termed this threshold as the 'lower threshold' of separation. Its complement, a 'higher threshold' also exists, above which the illusion can't be seen. From the experimental results, as illustrated in Figure 7.8a & 7.8b, we can see that for any subject the value of 'lower threshold' remains almost constant as the square size in Hermann grid changes. But the 'higher threshold' increases almost linearly as the size of the squares increases in the grid.



(a) (b)

Figure 7.8: (a) Graph of lower separation threshold vs. side length of Hermann grid squares. (b) Graph of higher separation threshold vs. side length of Hermann grid squares.

This once again is a type of scaling behavior as seen above in the case of Mach band illusion and Sine/Square grating illusions. This proportional variation of illusion size with stimulus size implies that the mechanism within the visual system that is generating this illusion is capable of generating illusions of a very wide range of length scales, from very small to very large. Any filtering model of brightness illusion will find it hard to reproduce such a wide range of length scales, since most image filters will have some typical

length scales and would not respond for stimuli much smaller or much larger than those typical length scale. The design of perceptual models of visual illusions should take such scaling behavior into consideration.

7.5 Abrupt disappearance of Hermann grid illusory blobs by tiny perturbations to the Hermann grid stimulus

In the previous section we saw how the Hermann grid illusion disappears when the grid square separation increases beyond a limit. Another way to disappear the Hermann grid illusion is by distorting the grid square shape into some other shape. In Figure 7.9a-7.9f we show three examples, from preexisting research [Lingelbach et al., 1985] [Spillmann, 1994] [Schiller and Carvey, 2005] [Geier, 2008], of various ways the grid may be distorted so that that the illusory blobs disappear. Geier et al. [Geier, 2008] had shown that the condition for complete wiping out of the illusion is to depart from the straight edges of the grids.

In this section we describe a novel modification of the Hermann grid stimulus in which tiny perturbations were added to the four corners of each of the grid squares, as a result of which the illusory blobs disappeared [Bakshi and Ghosh, 2020]. The size of perturbing squares was so small that overall only less than 2% of the total pixels of the original Hermann grid stimulus were changed. Even this small change was able to completely wipe out the illusory effect. Next we describe an experiment in which the modified version of the Hermann grid illusion was used as the stimulus, as shown in Figure 7.9h. The grid shown to our subjects contained 81 black squares in it. The

separation between the squares was kept fixed at 0.01 radians. This value of separation and the range of the square sizes were selected using the observations of the previous experiment described in section 7.4 so that the subjects could perceive the illusory blobs most prominently. The subjects were shown a modified Hermann grid, in which four tiny squares were added to each of the four corners of the Hermann grid squares. The added squares constitute less than 2% of the area of the unmodified Hermann grid stimulus. A series of such stimuli were shown one by one to the subjects with increasing width of the grid squares. Since the total number of squares was kept constant the size of the entire stimulus was also increasing accordingly. For each stimulus we changed the gray-level intensity of the added squares in steps, from white to black, as shown in Figures 7.10a, 7.10c, 7.10e, 7.10f, and the subjects were asked if the illusion was visible or not. As the gray-level intensities were varied, if any alteration in subject response occurred we recorded that particular gray-level intensity and from this information, we later calculated the gray-level intensity threshold of those tiny squares for a Hermann grid with specific square width. This procedure was repeated for different choices of square width. The largest grid in our experiments had squares of 0.04 radians width and the smallest one had 0.03 radians.

The most interesting outcome of this experiment is to show that very little modification of the Hermann grid (like adding very tiny squares with less than 2% area of the unmodified stimulus) is sufficient to wipe out the illusory perception of the dark blobs. Usage of the phrase "very little modification" will be more transparent if we compare our amount of

alterations in the Hermann grid with those in the existing literature [Figure 5 of Geier, 2008]. Gray-level intensity threshold vs. square side-length plots for the two subjects have been illustrated in Figure 7.11. From these graphs, it can be clearly observed that, for a particular subject, irrespective of the size of the squares, there is a fixed gray-level intensity threshold of the added perturbation below which the illusion becomes imperceptible. The size of the added squares doesn't need to be modified when the size of the stimulus is changed. Exactly same size of added squares could remove the illusory effect from the smallest as well as the biggest version of Hermann grid stimulus.

The modified Hermann grid stimulus shows that even tiny changes to a visual stimulus can drastically change the illusory effect. This implies that the visual mechanism producing this illusion can be extremely sensitive to the fine structure of the input stimulus. This observation would be very difficult to explain using spatial filtering models, since in these models tiny changes in input will only produce tiny changes to output and will not be able to produce such abrupt wipe out of the illusion as demonstrated by the modified Hermann grid stimulus.

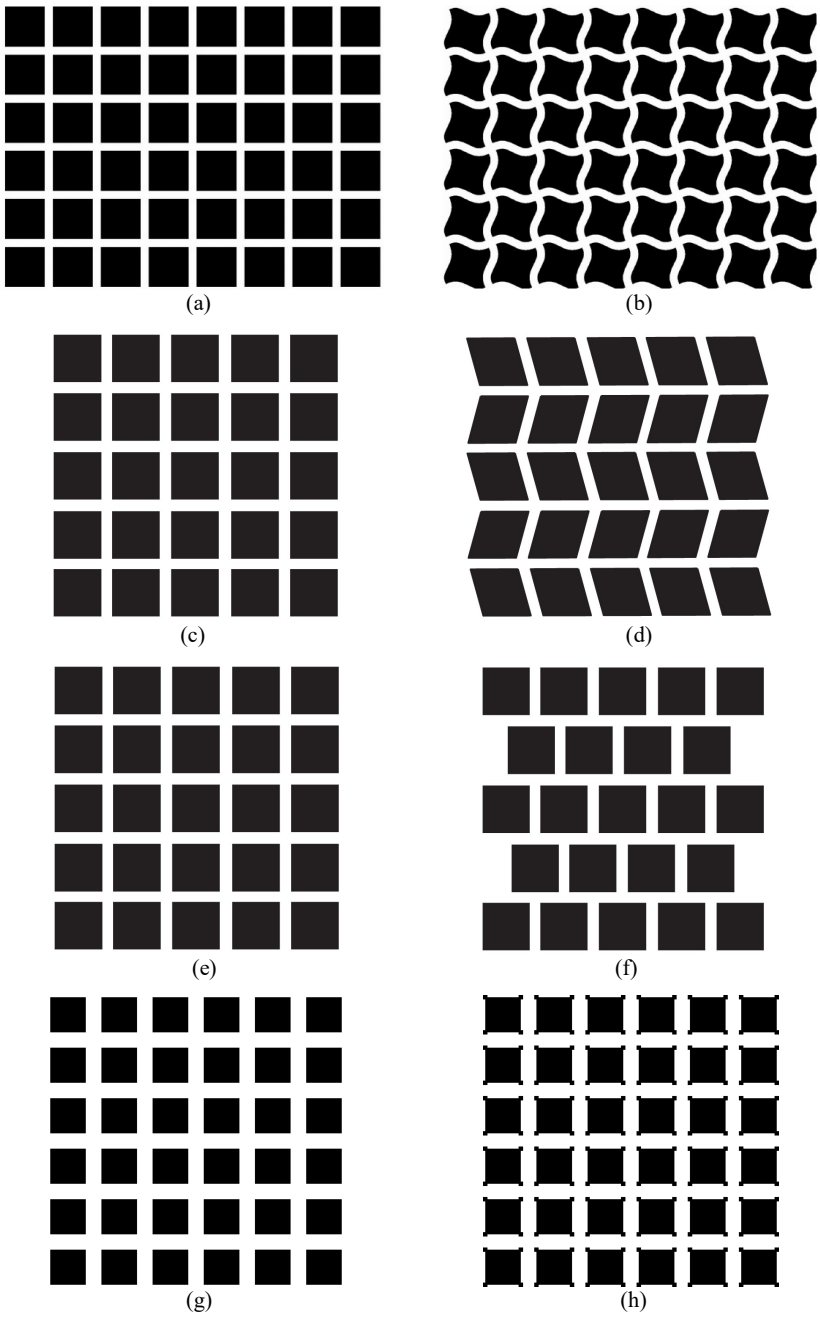
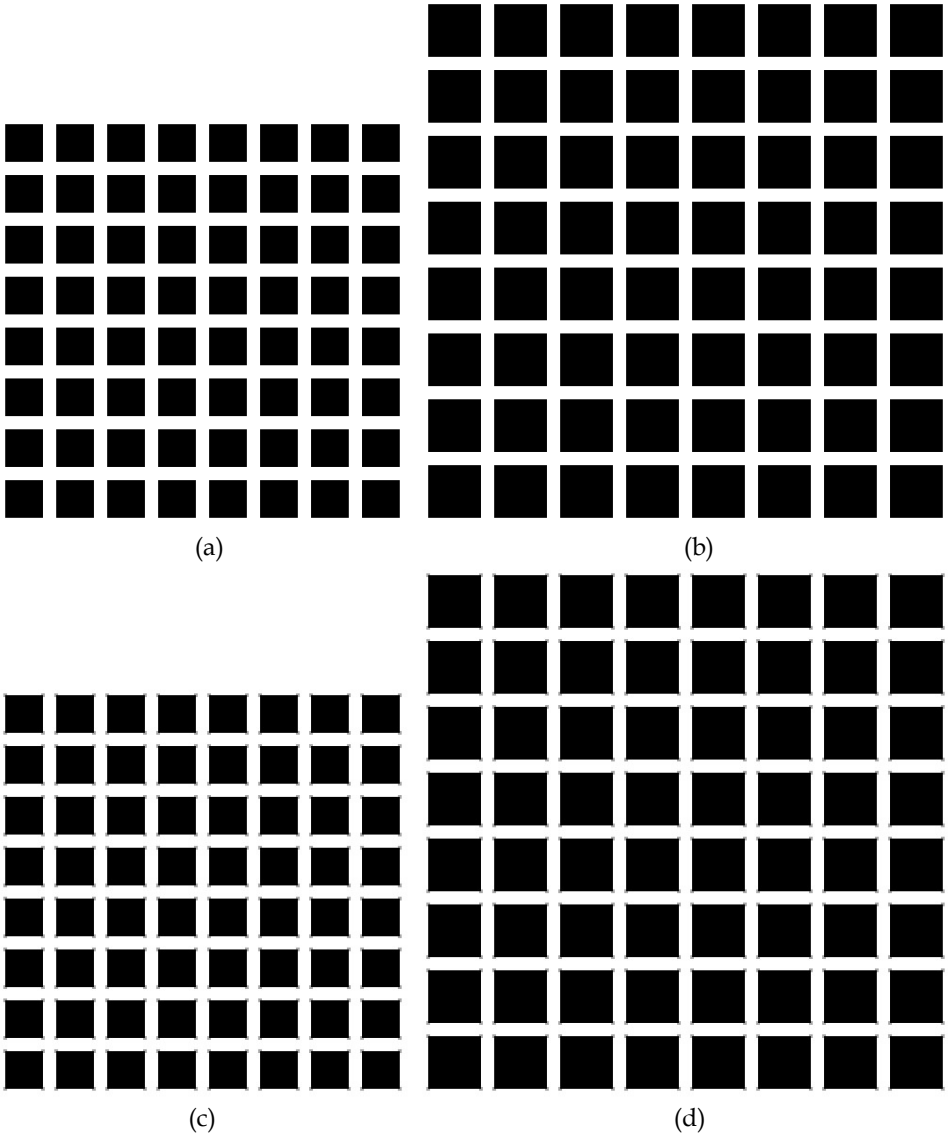


Figure 7.9: (a, c, e, g) The classical Hermann grid. (b) The distorted (sinusoid) grid reproduced from Geier et al. (2008). (d) The Schiller-Carvey (2005) modification with the vertical lanes distorted into a zigzag path (f) The Spillman (1994) modification where the relative positions of some of the grid squares have been changed (h) The proposed Hermann grid modification where the tiny perturbing squares are placed overlapping with each grid square corner. While modifications in (d) and (f) succeed in reducing the illusory strength, those in (b) and (h) completely wipes out the illusory spots.



(a) (b)

(c) (d)

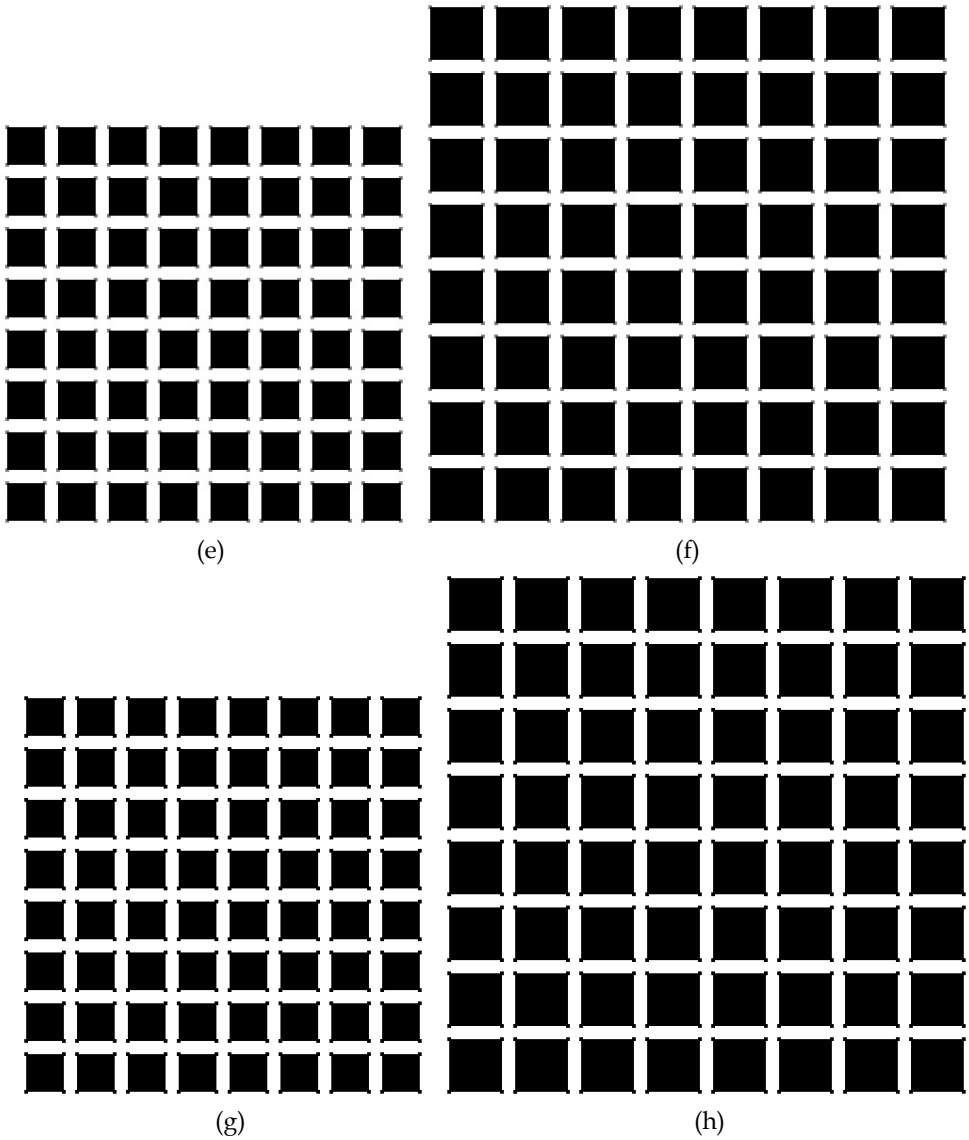


Figure 7.10: Modified Hermann grid stimulus with tiny gray squares added to the corners. The columns correspond to the size of the grid squares. The rows correspond to the intensity values of the tiny squares added to the corners of the large squares. The images on left column, (Figures a, c, e, g) correspond to small square sizes whereas the images on the right column (b, d, f, h) correspond to large square sizes. The gap between the grid squares is kept fixed. The first row (a, b) has tiny squares added to the corners with a grayscale value identical to the white background. The second row (c, d) has tiny squares with a high grayscale value (but less than in a, b) added to the corners. The tiny squares added to the third row (e, f) are darker than those on the second row. The tiny squares on the fourth row (g, h) have grayscale values same as the large squares of the Hermann grid.

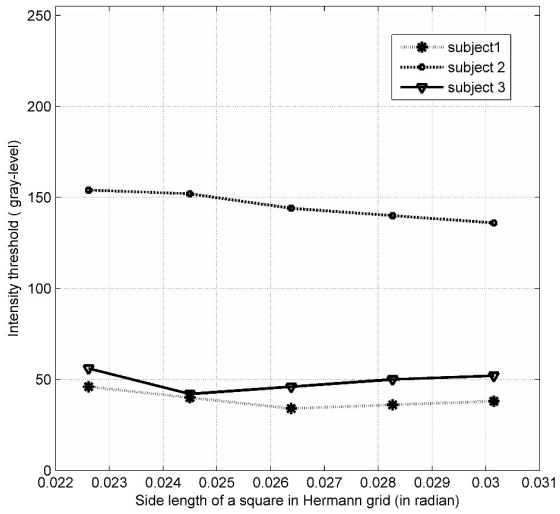


Figure 7.11: Graph of Intensity threshold vs. Side length of Hermann grid squares depicts that the gray level intensity threshold remains almost the same for each subject over a range of grid square length, with the size of the perturbing squares covering only about 1% of the original squares.

7.6 Concluding remarks on the problems of spatial filtering models and the way ahead

In the previous sections we demonstrated several situations which would be difficult to explain using spatial filtering algorithms. First we showed how several different types of illusions follow a scaling behavior in which the length scale of the illusory effect scales in proportion with the length scale of the input stimulus, from very small length scales to very large ones. This seemingly simple observation goes completely against the lateral inhibition based explanation of brightness illusions. Lateral inhibition, no matter how it occurs biologically, will only produce responses of a finite range of length scales, unless we assume that the size of receptive field itself is variable and can change its size from extremely small to extremely large. There have only occasionally been some such attempts towards building multiscale models of visual illusions [Blakeslee and McCourt, 1997]

[Nematzadeh et al., 2017]. Any moderately sized response function will neither respond to signals of very small scale nor signals of very large scale [Bakshi and Ghosh, 2012a] [Bakshi and Ghosh, 2012b]. Secondly we showed how even a very tiny change to the visual stimulus is capable of completely destroying the illusory effect. This sensitivity to tiny changes in the input stimulus will also be very difficult to explain with spatial filtering algorithms since most spatial filtering algorithms have a continuous nature, which means that a small change to the input will produce only a small change to the output and therefore will not be able to wipe out the illusory effect completely [Bakshi and Ghosh, 2020].

Several approaches suggest themselves in trying to solve the above two problems. For example, fractal based spatial filtering approaches may be used to produce responses at all possible length scales. Another approach could be to use wavelet based filters which show the property of scale invariance. The second problem of sensitivity to input signal may be solved with a nonlinear dynamical system approach which also shows the property of extreme sensitivity to initial conditions.

Spehar et al. have reported before [Spehar et al., 2003], [Spehar et al., 2015] that humans show an aesthetic preference for visual patterns having a fractal nature because images of natural scenes show a similar fractal nature. This aesthetic preference is correlated with visual sensitivity, i.e. the ability to distinguish between two slightly differentiated patterns. The human visual system shows higher visual sensitivity for patterns having a fractal nature than patterns of non-fractal nature. This ability to distinguish between

patterns is high for fractal patterns with a certain range of fractal dimensions and low for patterns outside this range or for non-fractal patterns. This suggests that the pattern recognition ability of the visual system is different for different fractal or non-fractal patterns and there must exist regions within the visual system for recognition of these fractal patterns. Combining this fractal based approach with a multiresolution representation of 2D signals, such as the wavelet transform, could help in building useful models which exhibit this property of scale invariance [Field, 1993] [Field, 1999] [Van De Ville and Unser, 2008].

A non-linear dynamical model could be used to model the property of sensitivity to small changes in the input stimulus [Ditzinger, 2010] [Hock and Schöner, 2016] [Richards et al., 1994]. In such a model the output would be computed by recursively applying a non-linear transformation to the input stimulus which would then iteratively converge to the final output. If the non-linear transformation is chosen carefully it is possible that it may exhibit sensitivity to initial conditions which would produce drastic changes in the final output, like the example shown in the Hermann Grid illusion in this chapter [Bakshi and Ghosh, 2020].

Exploring such approaches to solve the scale problem and the sensitivity problem should be a priority in future.

Appendix A

A.1 Detailed experimental procedure for the comparison of performance of M-channel vs. Bilateral Filter in implementing "vision at a glance".

The main components of the experimental setup include a standard set of images, a Gaussian noise adding system, algorithms that implement the Bilateral and DDOG filters and a set of parameters to compare the performance of the filters implementing 'vision at a glance' with respect to the original (noise free) images. The workflow is depicted in the following flowchart. The detailed methodology is stated in the following subsections.

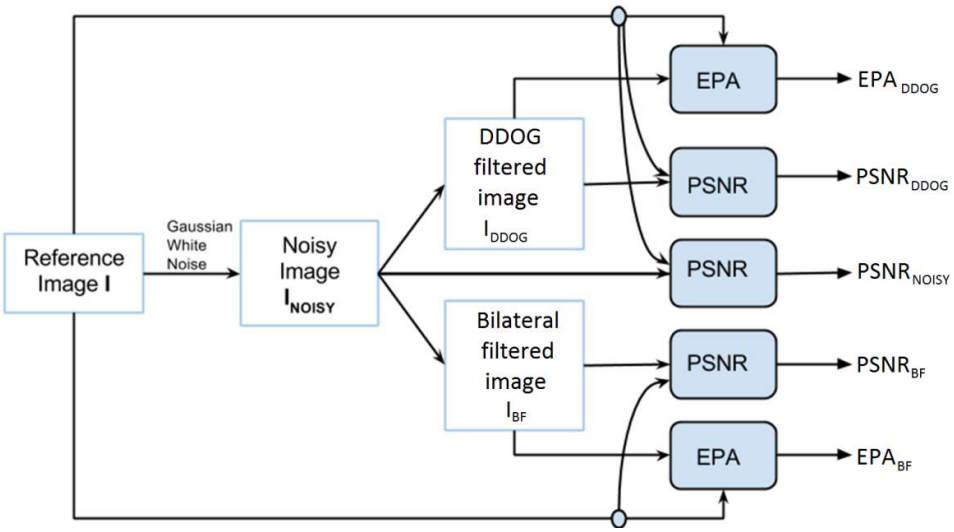


Figure A.1: Schematic of comparison of the DDOG based M-channel filter with the bilateral filter.

A.1.1 Set of images (S)

Since, the natural 'vision at a glance' is unlikely to provide us with any detail about the color information of an image because of the supposedly poor chromatic selectivity of the M channel [Merigan et al., 1991] [Merigan

and Maunsell, 1993], we are motivated to use gray-scale images for the experiment with both the natural (M) filter and the artificial (bilateral) filter. We use a standard set of 26 gray scale images [Heath et al., 1998] [http://marathon.csee.usf.edu/edge/edge_detection.html] as reference images for our experiment. This reference set consists of both indoor and outdoor scenes and both natural and manmade objects.

A.1.2 The Experiment

For a given reference image (I) from our set of images(S), we add zero-mean, Gaussian white noise of local variance v, to obtain a noisy image (INOISY). Then we apply standard DDOG based M-channel filter on INOISY to obtain filtered image IDDOG. We use spatial-domain standard deviation = 3 and the intensity-domain standard deviation = 0.1 for bilateral filter and then apply it on INOISY to obtain filtered image IBF.

We apply 5 different levels of zero-mean, Gaussian white noise of local variance v (v = 0.1 , 0.05 , 0.01 , 0.005 , 0.001) to each image and carry out the above mentioned experiment to demonstrate the efficiency of filtering in presence of varied levels of noise.

Thus after the completion of the experiment, corresponding to the reference image set, we have the noisy image set, the M-channel filtered image set and the bilateral filtered image set corresponding to each value of v.

A.1.3 Edge detector used for obtaining edge map

We have chosen the Canny edge detection method [Canny, 1986] for obtaining the edge map of the reference image (I) and the edge maps of the

corresponding filtered images. The Canny method finds edges by looking for local maxima of the gradient of I. The gradient is calculated using the derivative of a Gaussian filter. The method uses two thresholds, to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges. We have used the Matlab function **edge** [http://www.mathworks.in/help/images/ref/edge.html] with the parameters set as follows: thresh = 0.1, sigma = 0.975 and type = 'canny'.

A.1.4 Parameters to compare the filter performances in implementing 'vision at a glance'

The primary objective for vision at a glance is to smoothen the images to reduce fine level details while preserving the edge information. This objective needs to be achieved without much computational complexity, since we are focused on 'vision at a glance' and not 'vision with scrutiny' [Hochstein and Ahissar, 2002]. We propose to use two metrics for comparing how well the filters achieve our objective.

a) **Peak Signal to Noise Ratio(PSNR)**

To compute the PSNR, we first calculate the mean-squared error (MSE) using the following equation:

$$\boldsymbol{MSE} = \frac{\sum_{M,N}[\boldsymbol{I_1}(\boldsymbol{m},\boldsymbol{n})-\boldsymbol{I_2}(\boldsymbol{m},\boldsymbol{n})]^2}{\boldsymbol{M*N}} \qquad \textbf{(A.1)}$$

I_1 is the reference image and I_2 is the image of which PSNR is being calculated. M and N are the number of rows and columns in the input images, respectively. Then the PSNR is calculated using the following equation.

$$PSNR = 10log_{10}\left(\frac{R^2}{MSE}\right) \quad \textbf{(A.2)}$$

where R is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255.

b) Edge preservation accuracy (EPA)

For a given image **I**, and corresponding reference $\mathbf{I_{REF}}$, with corresponding edge map E and E_{REF} respectively, a pixel in an edge map is classified into an edge pixel if it is part of an edge, else it is classified as a non edge pixel.

EPA [Heath et al., 1998] [Bowyer et al., 2001] [Liu and Haralick, 2000] is defined as follows:

$$EPA = \frac{TP+TN}{P} \quad \textbf{(A.3)}$$

Where P is the total number of pixels in E. TP is the count of edge pixels present at E(x, y), and also present at E_{REF} (x,y) and TN is the count of edge pixels absent at E(x, y), and also absent at E_{REF} (x,y) .

PSNR helps in providing us with a measure of the extent of smoothing achieved by filtering, while EPA provides a measure of the extent of edge information preserved post filtering.

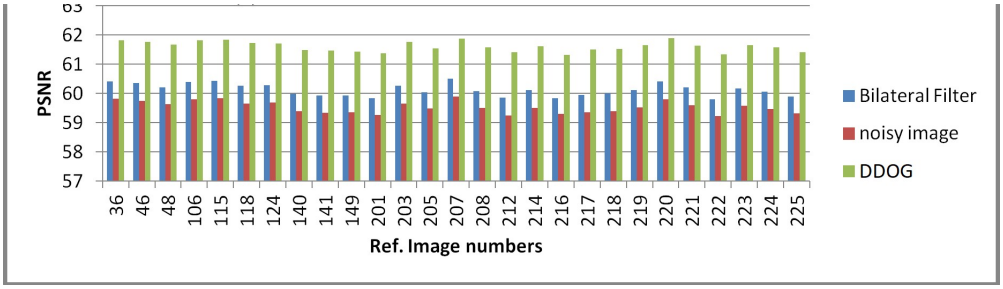
A.1.5 Analysis of the filtered images

The PSNR of the filtered images (I_{DDOG} and I_{BF}) is compared with that of I_{NOISY}. The PSNR in each of the cases is measured with respect to a reference image (I). We apply the Canny edge detector on I, I_{DDOG} and I_{BF} with the parameters mentioned earlier. We obtain 3 edge maps say E, E_{DDOG} and E_{BF} respectively. Each of this edge maps are black and white images with white

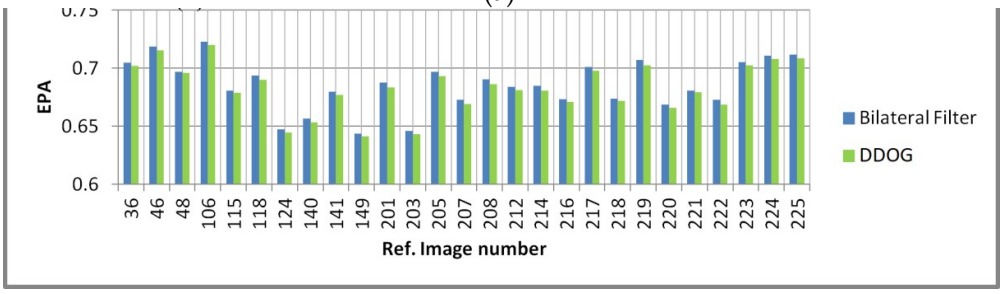
regions marking out the edges. Using E as reference edge map, we calculate edge preservation accuracy EPA_{DDOG} and EPA_{BF} corresponding to E_{DDOG} and E_{BF} respectively.

A.2 Results and discussion

The parameters defined in section A.1.4 are used to compare the performance of the two concerned filters at various noise levels. These have been presented in Figures A.2-A.6.

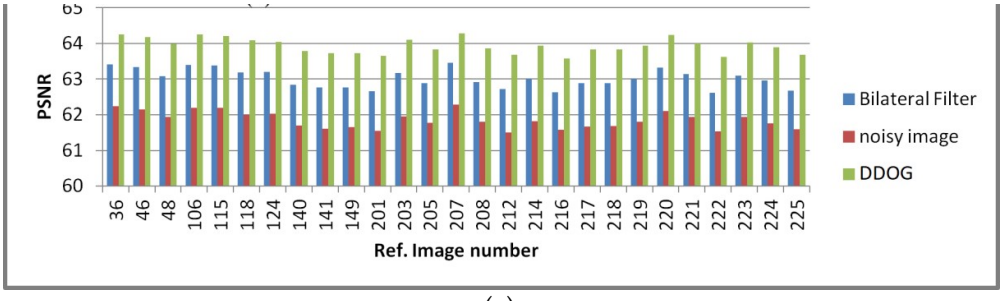


(a)

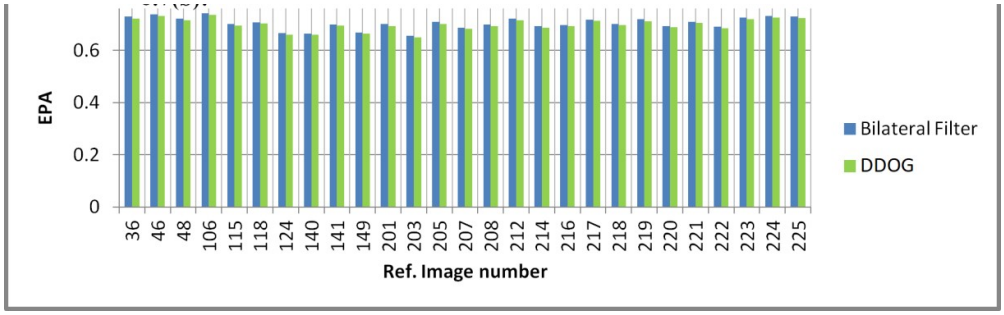


(b)

Figure A.2 (a) PSNR comparison (v=0.1 of 'gaussian noise') (b) Edge Preservation Accuracy comparison (variance = 0.1)

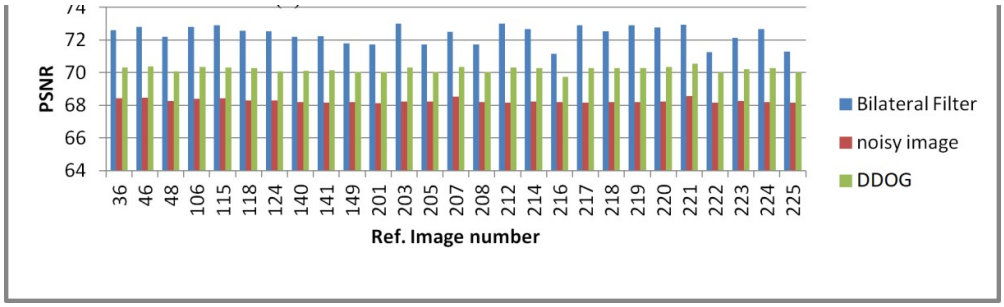


(a)

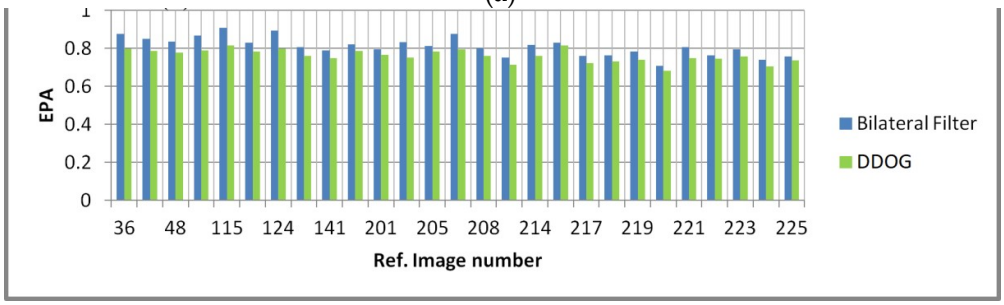


(b)

Figure A.3 (a) PSNR comparison (variance=0.05) (b) Edge Preservation Accuracy comparison (variance = 0.05)

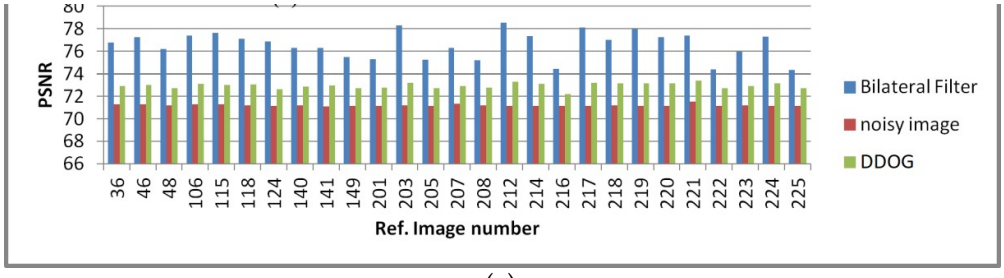


(a)

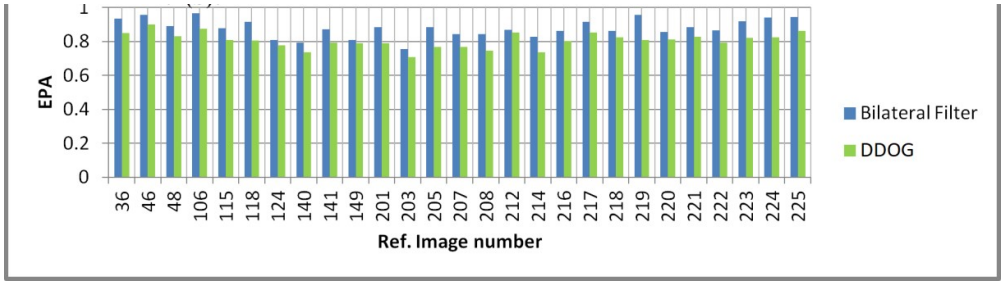


(b)

Figure A.4 (a) PSNR comparison (variance=0.01) (b) Edge Preservation Accuracy comparison (variance = 0.01)

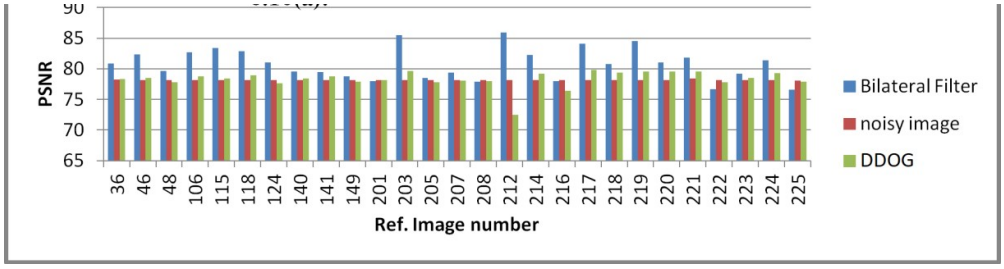


(a)

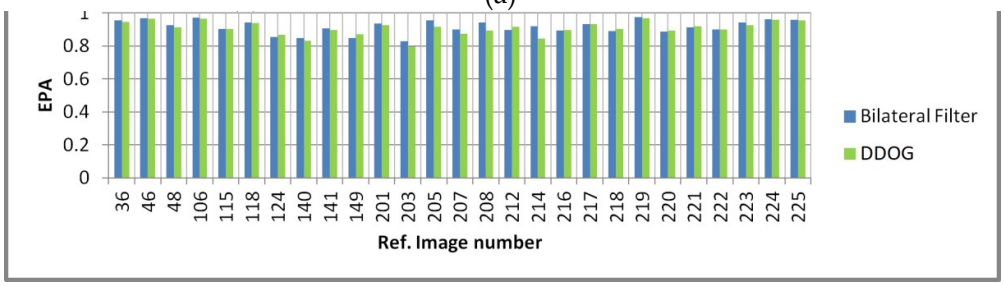


(b)

Figure A.5 (a) PSNR comparison (variance=0.005) (b) Edge Preservation Accuracy comparison (variance = 0.005)



(a)



(b)

Figure A.6 (a) PSNR comparison (variance=0.001) (b) Edge Preservation Accuracy comparison (variance = 0.001)

Analyzing the results demonstrated in Figures A.2-A.6, and comparing the performance of the bilateral filter vis-a-vis the M DDOG filter in image smoothing and edge preservation, one can draw the following conclusions:

1) For images containing high level of noise (i.e. white Gaussian noise having high variance with v = 0.1 and 0.05), PSNR of M-channel filtered image is substantially and consistently greater than the PSNR of bilateral filtered image. On the other hand, in case of very low noise images (i.e. noise having low variance v =

0.01, 0.005 and 0.001), the bilateral filter performs better than the M channel filter in terms of PSNR.

2) For all values of variances of white Gaussian noise, the EPA for bilateral filters is only marginally better than the corresponding M-channel filtered images and both may be considered equal for all practical purposes.

3) Most significantly, for 512x512 size gray scale images, M-channel filtering runs 193 times faster on average, than bilateral filtering on our machine. Since we have a linear combination of Gaussian functions for M-channel filter, the kernel of filtering can be pre-computed which helps in achieving faster implementation. On the other hand, the kernel of filtering for bilateral filter needs to be computed at runtime since it varies according to the input image. Hence M-channel filtering is significantly faster than bilateral filtering.

For each pixel p in a m x m mask, bilateral filter has to compute the difference of every pixel intensity (I_p) with the mask's center pixel intensity (I_c) and thereafter use this difference to compute the Gaussian $\mathbf{G(|I_p - I_c|)}$. In case of M-channel filter, this computation is not required.

We can formally state it as follows:

Let the size of filter be m x m and the size of the image be n x n.

The number of times the filter has to be applied on the image = n x n.

The number of extra computations per computation of bilateral filter kernel = m x m

Thus the total number of extra computation for the entire image = (m x m) x n x n

The plot in Figure A.7 shows how the number of extra computations that bilateral filters performs increases with increase in size of image for a filter size of 3x3.

4) Thus from points 1), 2) & 3) we may claim that the performance of both M-channel and Bilateral filters in removing noise is substantial. Moreover with the M-channel filter, we achieve higher PSNR as compared to the noisy images and at the same time, we have achieved an EPA of at least 0.6 for all values of variance of white Gaussian noise added to our reference images. So with the specified settings of bilateral filter and DDOG based M-channel filter we have been able to remove noise while preserving edge details. Although both the filters are appropriate in achieving 'vision at a glance' akin to image preprocessing in computer vision, the M-channel filter shows better overall performance for high noise images and acceptable performance for very low noise images while maintaining much higher computational performance all throughout, and is therefore highly appropriate for real time applications.

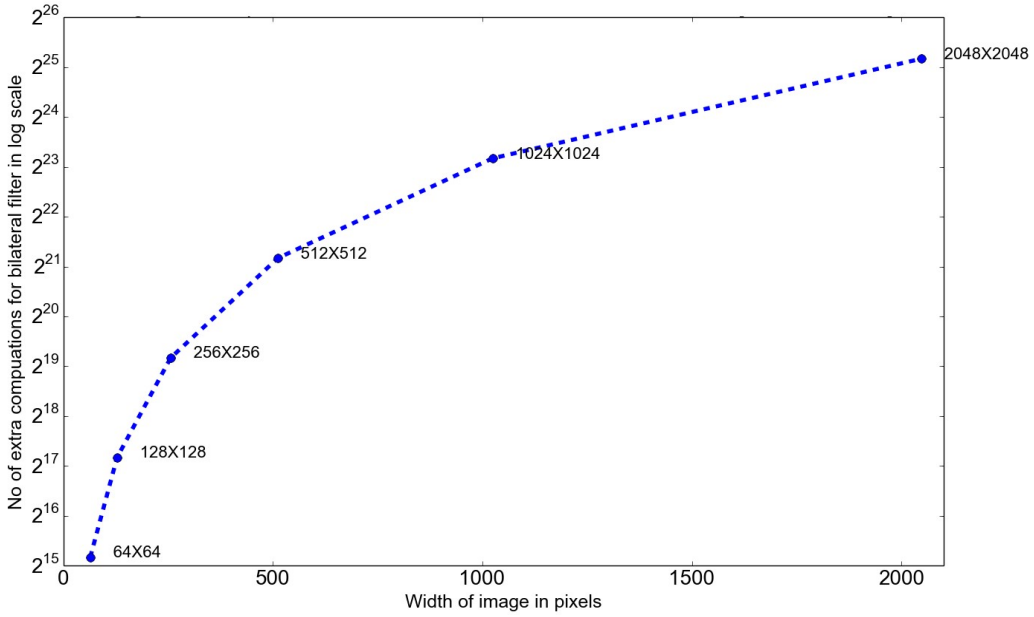


Figure A.7 Computation overhead of Bilateral filter with respect to the M-channel filter [Filter size = 3 x 3]

Bibliography

[Bakshi and Ghosh, 2012a] Bakshi A., & Ghosh K. (2012). Scaling Properties of Mach Bands and Perceptual Models. In: Kundu M.K., Mitra S., Mazumdar D., Pal S.K. (eds) *Perception and Machine Intelligence. PerMIn 2012.* Lecture Notes in Computer Science, vol 7143. Springer, Berlin, Heidelberg.

[Bakshi and Ghosh, 2012b] Bakshi, A. & Ghosh, K. (2012). Some insights into why the perception of mach bands is strong for luminance ramps and weak or vanishing for luminance steps. *Perception*, 41(11):1403–1408.

[Bakshi and Ghosh, 2015] Bakshi A., & Ghosh K. (2015). Perceiving and Modeling Brightness Contradictions Through the Study of Brightness Illusions. In: Beziau JY., Chakraborty M., Dutta S. (eds) *New Directions in Paraconsistent Logic.* Springer Proceedings in Mathematics & Statistics, vol 152 (pp. 447-463). Springer, New Delhi.

[Bakshi et al., 2016] Bakshi, A., Roy, S., Mallick, A., & Ghosh, K. (2016). Limitations of the oriented difference of gaussian filter in special cases of brightness perception illusions. *Perception*, 45(3):328-336.

[Bakshi and Ghosh, 2017] Bakshi, A., & Ghosh, K. (2017). A Neural Model of Attention and Feedback for Computing Perceived Brightness in Vision. In: *Handbook of Neural Computation* (pp. 487-513). Academic Press.

[Bakshi et al., 2017] Bakshi, A., Roy, S., Mallick, A., & Ghosh, K. (2017) Experimental Observations of the Visibility Threshold of Illusory effects in Hermann Grid, Sinusoidal and Square Gratings and their Possible Implications. *Int J Open Access Ophthal* 2(1):1-8. DOI: 10.15226/2474-9249/2/1/00120

[Bakshi and Ghosh, 2018] Bakshi, A. & Ghosh, K. (2018). A parsimonious model of brightness induction. *Biological Cybernetics*, 112(3):237-251.

[Bakshi and Ghosh, 2020] Bakshi, A., & Ghosh, K. (2020). Tiny Squares at the Hermann Grid Corners Can Completely Remove the Illusion. *Perception*, 49(2): 232-239.

[Bakshi et al., 2021] Bakshi, A., Roy, S., Mallick, A., & Ghosh, K. (2021). A discrete magno–parvo additive model in early vision for explaining brightness perception in varying contrastive contexts. *Biological Cybernetics*, (published online: 1-17). https://doi.org/10.1007/s00422-021-00896-4

[Bar, 2003] Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, 15(4):600–609.

[Békésy, 1967] von Békésy, G. (1967). Mach band type lateral inhibition in different sense organs. *The Journal of general physiology, 50*(3), 519-532.

[Békésy, 1968] von Bekesy, G. (1968). Brightness distribution across the Mach bands measured with flicker photometry, and the linearity of sensory nervous interaction. *JOSA, 58*(1), 1-8.

[Bhattacharjee et al., 2015] Bhattacharjee, D., Bakshi, A., & Ghosh, K. (2015). Comparison between an HVS inspired linear filter and the bilateral filter in performing "vision at a glance" through smoothing with edge preservation. *International Journal of Image and Graphics*, 15(04):1550015.

[Blakeslee and McCourt, 1997] Blakeslee, B. & McCourt, M. E. (1997). Similar mechanisms underlie simultaneous brightness contrast and grating induction. *Vision Research*, 37(20):2849–2869.

[Blakeslee and McCourt, 1999] Blakeslee, B. & McCourt, M. E. (1999). A multiscale spatial filtering account of the white effect, simultaneous brightness contrast and grating induction. *Vision Research*, 39(26):4361–4377.

[Blakeslee and McCourt, 2004] Blakeslee, B. & McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research*, 44(21):2483–2503.

[Blakeslee et al., 2005] Blakeslee, B., Pasieka, W., & McCourt, M. E. (2005). Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Research*, 45(5):607–615.

[Blakeslee et al., 2016] Blakeslee, B., Cope, D., & McCourt, M. E. (2016). The Oriented Difference of Gaussians (ODOG) model of brightness perception: Overview and executable Mathematica notebooks. *Behavior research methods*, 48(1):306-312.

[Bowyer et al., 2001] Bowyer, K., Kranenburg, C., & Dougherty, S. (2001). Edge detector evaluation using empirical ROC curves. *Computer Vision and Image Understanding, 84*(1), 77-103.

[Bullier, 2001] Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, 36(2-3):96–107.

[Canny, 1986] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6), 679-698.

[Cao et al., 2015] Cao, C., Liu, X., Yang, Y., Yu, Y., Wang, J., Wang, Z., Huang, Y., Wang, L., Huang, C., Xu, W., et al. (2015). Look and think twice: Capturing top-down visual attention with feedback convolutional neural networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2956–2964.

[Carrasco et al., 2004] Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, 7(3):308.

[Croner and Kaplan, 1995] Croner, L. J. & Kaplan, E. (1995). Receptive fields of p and m ganglion cells across the primate retina. *Vision Research*, 35(1):7–24.

[Dade] Dade, K. Toonify: Cartoon Photo Effect Application.

[Dasgupta et al., 2013] Dasgupta, A., Bakshi, A., & Ghosh, K. (2013, February). Lateral inhibition based holistic approach to adaptive image enhancement. In *2013 3rd IEEE International Advance Computing Conference (IACC)* (pp. 1143-1148). IEEE.

[Davies and Morland, 2002] Davies, N. P. & Morland, A. B. (2002). The Hermann-Hering grid illusion demonstrates disruption of lateral inhibition processing in diabetes mellitus. *British Journal of Ophthalmology*, 86(2): 203-208.

[De Monasterio and Gouras, 1975] De Monasterio, F. & Gouras, P. (1975). Functional properties of ganglion cells of the rhesus monkey retina. *The Journal of Physiology*, 251(1):167–195.

[De Valois and De Valois, 1988] De Valois, R. & De Valois, K. (1988). Oxford psychology series, no. 14. *Spatial Vision*. New York, NY, US: Oxford University Press.

[De Valois et al., 2000] De Valois, R. L., Cottaris, N. P., Mahon, L. E., Elfar, S. D., & Wilson, J. A. (2000). Spatial and temporal receptive fields of geniculate and cortical cells and directional selectivity. *Vision research, 40*(27), 3685-3702.

[Ditzinger, 2010] Ditzinger, T. (2010). Optical illusions: examples for nonlinear dynamics in perception. In *Nonlinear Dynamics in Human Behavior* (pp. 179-191). Springer, Berlin, Heidelberg.

[Durand and Dorsey, 2002] Durand, F., & Dorsey, J. (2002, July). Fast bilateral filtering for the display of high-dynamic-range images. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (pp. 257-266).

[Ehrenstein and Ehrenstein, 1999] Ehrenstein, W. H. & Ehrenstein, A. (1999). Psychophysical methods. *Modern Techniques in Neuroscience Research*, page 1211-1241.

[Fechner, 1860] Fechner, G. (1860). *Elemente der Psychophysik*. Number v. 1 in Elemente der Psychophysik. Breitkopf und Härtel.

[Ferrera et al., 1992] Ferrera, V. P., Nealey, T. A., & Maunsell, J. H. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area v4. *Nature*, 358(6389):756.

[Field, 1993] Field, D. J. (1993). Scale-invariance and Self-similar 'Wavelet' Transforms: an Analysis of Natural Scenes and Mammalian Visual Systems. In: *Wavelets, Fractals and Fourier Transforms: New Developments and New Applications.* Farge, M., Hunt, J. and Vassilicos., eds. Oxford University Press. (1993)

[Field, 1999] Field, D. J. (1999). Wavelets, vision and the statistics of natural scenes. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1760):2527-2542.

[Fiorentini et al., 1990] Fiorentini, A., Baumgartner, G., Magnussen, S., Schiller, P. H., & Thomas, J. P. (1990). *The perception of brightness and darkness: Relations to neuronal receptive fields.*

[Foley and McCourt, 1985] Foley, J. M. & McCourt, M. E. (1985). Visual grating induction. *JOSA A, 2*(7):1220–1230.

[Fomm, 1896] Fomm, L. (1896). The wavelength of roentgen-rays. *Annalen der Physik, 59*:350–353.

[Geier, 2008] Geier, J., Bernáth, L., Hudák, M., & Séra, L. (2008). Straightness as the main factor of the Hermann grid illusion. *Perception, 37*(5), 651-665.

[Gescheider, 1997] Gescheider, G. A. (1997). *Psychophysics: the fundamentals.* L. Erlbaum Associates.

[Ghosh, 2012] Ghosh, K. (2012). A possible role and basis of visual pathway selection in brightness induction. *Seeing and Perceiving*, 25(2):179–212.

[Ghosh, 2017] Ghosh, K. (2017). Experimental observations of the visibility threshold of illusory effects in hermann grid, sinusoidal and square gratings and their possible implications. *International Journal of open Access Ophthalmology*, 2(1):18.

[Ghosh and Pal, 2010] Ghosh, K. & Pal, S. K. (2010). Some insights into brightness perception of images in the light of a new computational model of figure–ground segregation. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(4):758–766.

[Ghosh et al., 2005a] Ghosh, K., Sarkar, S., & Bhaumik, K. (2005, January). Low-level brightness-contrast illusions and non classical receptive field of mammalian retina. In *Proceedings of 2005 International Conference on Intelligent Sensing and Information Processing, 2005*. (pp. 529-534). IEEE.

[Ghosh et al., 2005b] Ghosh, K., Sarkar, S., & Bhaumik, K. (2005b). A possible mechanism of zero-crossing detection using the concept of the extended classical receptive field of retinal ganglion cells. *Biological Cybernetics*, 93(1):1–5.

[Ghosh et al., 2006] Ghosh, K., Sarkar, S., & Bhaumik, K. (2006). A possible explanation of the low-level brightness–contrast illusions in the light of an extended classical receptive field model of retinal ganglion cells. *Biological Cybernetics*, 94(2):89–96.

[Gilchrist et al., 1999] Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., Spehar, B., Annan, V., & Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, 106(4):795.

[Gilchrist, 2006] Gilchrist, A. (2006). *Seeing black and white* (No. 40). OUP USA.

[Gilchrist, 2014] Gilchrist, A. (2014). A gestalt account of lightness illusions. *Perception*, 43(9):881–895.

[Gonzalez and Woods, 2003] Gonzalez, R. C. & Woods, R. E. (2003). *Digital Image Processing*. Pearson.

[Grossberg et al., 1997] Grossberg, S., Mingolla, E., & Ross, W. D. (1997). Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neurosciences*, 20(3):106–111.

[Grossberg and Todorovic, 1988] Grossberg, S. & Todorovic, D. (1988). Neural dynamics of 1-d and 2-d brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, 43(3):241–277.

[Heath et al., 1998] 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.

[Heinemann, 1955] Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing-and test-field luminances. *Journal of Experimental Psychology*, 50(2):89.

[Hermann, 1870] Hermann, L. (1870). Eine erscheinung simultanen contrastes. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, 3(1):13–15.

[Hochstein and Ahissar, 2002] Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36*(5), 791-804.

[Hock and Schöner, 2016] Hock, H. S., & Schöner, G. (2016). Nonlinear dynamics in the perceptual grouping of connected surfaces. *Vision research, 126*:80-96.

[Howe, 2001] Howe, P. D. (2001). A comment on the Anderson (1997), the Todorović (1997), and the Ross and Pessoa (2000) explanations of White's effect. *Perception, 30*(8), 1023-1026.

[Hubel and Wiesel, 1959] Hubel, D. H. & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3):574–591.

[Hubel and Wiesel, 1962] 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):106–154.

[Hubel and Wiesel, 1968] Hubel, D. H. & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1):215–243.

[Hupé et al., 1998] Hupé, J., James, A., Payne, B., Lomber, S., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by v1, v2 and v3 neurons. *Nature*, 394(6695):784.

[Kandel et al., 2000] Kandel, E. R., Schwartz, J. H., & Jessell, T. M. of Biochemistry, D., Jes-sell, MBT, Siegelbaum, S., & Hudspeth, A. (2012). *Principles of Neural Science*.

[Kaplan and Shapley, 1982] Kaplan, E. & Shapley, R. (1982). X and y cells in the lateral geniculate nucleus of macaque monkeys. *The Journal of Physiology*, 330(1):125–143.

[Kingdom and Moulden, 1992] Kingdom, F. & Moulden, B. (1992). A multi-channel approach to brightness coding. *Vision Research*, 32(8):1565–1582.

[Kingdom, 2011] Kingdom, F. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, 51(7):652–673.

[Kveraga et al., 2007] Kveraga, K., Boshyan, J., & Bar, M. (2007). Magnocellular projections as the trigger of top-down facilitation in recognition. *Journal of Neuroscience*, 27(48):13232–13240.

[Land and McCann, 1971] Land, E. H. & McCann, J. J. (1971). Lightness and retinex theory. *JOSA*, 61(1):1–11.

[Lingelbach et al., 1985] Lingelbach, B., Block, B., Hatzky, B., & Reisinger, E. (1985). The Hermann grid illusion—Retinal or cortical? *Perception, 14(supplement, Abstract)*, A7.

[Liu and Haralick, 2000] Liu, G., & Haralick, R. M. (2000, June). Assignment problem in edge detection performance evaluation. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)* (Vol. 1, pp. 26-31). IEEE.

[Lv et al., 2016] Lv, Q., Wang, B., & Zhang, L. (2016). Saliency computation via whitened frequency band selection. *Cognitive Neurodynamics*, 10(3):255–267.

[Marr, 1982] Marr, D. (1982). *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information*. MIT Press.

[Maunsell et al., 1990] Maunsell, J., Nealey, T. A., & DePriest, D. D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (mt) of the macaque monkey. *Journal of Neuroscience*, 10(10):3323–3334.

[McAlonan et al., 2008] McAlonan, K., Cavanaugh, J., & Wurtz, R. H. (2008). Guarding the gateway to cortex with attention in visual thalamus. *Nature*, 456(7220):391.

[McCourt, 1982] McCourt, M. E. (1982). A spatial frequency dependent grating-induction effect. *Vision Research*, 22(1):119–134.

[McCourt and Foley, 1985] McCourt, M. E., & Foley, J. M. (1985). Spatial frequency interference on grating-induction. *Vision Research*, 25(10), 1507-1518.

[McCourt et al., 2016] McCourt, M. E., Blakeslee, B., & Cope, D. (2016). The oriented difference-of-Gaussians model of brightness perception. *Electronic Imaging*, 2016(6):1-9.

[Merigan and Maunsell, 1993] Merigan, W. H. & Maunsell, J. H. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience*, 16(1):369–402.

[Merigan et al., 1991] Merigan, W. H., Katz, L. M., & Maunsell, J. H. (1991). The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. *Journal of Neuroscience, 11*(4), 994-1001.

[Morrone et al., 1986] Morrone, M. C., Ross, J., Burr, D. C., & Owens, R. (1986). Mach bands are phase dependent. *Nature*, 324(6094):250–253.

[Nematzadeh et al., 2017] Nematzadeh, N., Powers, D. M., & Lewis, T. W. (2017). Bioplausible multiscale filtering in retino-cortical processing as a mechanism in perceptual grouping. *Brain informatics*, *4*(4): 271-293.

[Nowak et al., 1995] Nowak, L., Munk, M., Girard, P., & Bullier, J. (1995). Visual latencies in areas v1 and v2 of the macaque monkey. *Visual Neuroscience*, 12(2):371–384.

[Pessoa, 1996] Pessoa, L. (1996). Mach-band attenuation by adjacent stimuli: experiment and filling-in simulations. *Perception*, 25(4):425–442.

[Paris and Durand, 2006] Paris, S., & Durand, F. (2006, May). A fast approximation of the bilateral filter using a signal processing approach. In *European conference on computer vision* (pp. 568-580). Springer, Berlin, Heidelberg.

[Paris et al., 2009] Paris, S., Kornprobst, P., Tumblin, J., & Durand, F. (2009). *Bilateral filtering: Theory and applications.* Now Publishers Inc.

[Ratliff, 1965] Ratliff, F. (1965). *Mach bands: quantitative studies on neural networks*. HoldenDay, San Francisco London Amsterdam.

[Ratliff and Hartline, 1959] Ratliff, F. & Hartline, H. K. (1959). The responses of limulus optic nerve fibers to patterns of illumination on the receptor mosaic. *The Journal of General Physiology*, 42(6):1241–1255.

[Ratliff et al., 1983] Ratliff, F., Milkman, N., & Rennert, N. (1983). Attenuation of mach bands by adjacent stimuli. *Proceedings of the National Academy of Sciences*, 80(14):4554–4558.

[Richards et al., 1994] Richards, W., Wilson, H. R., & Sommer, M. A. (1994). Chaos in percepts?. *Biological cybernetics, 70*(4):345-349.

[Robinson et al., 2007a] Robinson, A. E., Hammon, P. S., & de Sa, V. R. (2007a). Explaining brightness illusions using spatial filtering and local response normalization. *Vision Research*, 47(12):1631–1644.

[Robinson et al., 2007b] Robinson, A. E., Hammon, P. S., & de Sa, V. R. (2007b). A filtering model of brightness perception using frequency-specific locally-normalized oriented difference-of gaussians (flodog). *Journal of Vision*, 7(9):237–237.

[Rodieck and Stone, 1965] Rodieck, R. W. & Stone, J. (1965). Analysis of receptive fields of cat retinal ganglion cells. *Journal of Neurophysiology*, 28(5):833–849.

[Schiller and Carvey, 2005] Schiller, P. H., & Carvey, C. E. (2005). The Hermann grid illusion revisited. *Perception, 34*(11), 1375-1397.

[Schiller and Malpeli, 1978] Schiller, P. H. & Malpeli, J. G. (1978). Functional specificity of lateral geniculate nucleus laminae of the rhesus monkey. *Journal of Neurophysiology*, 41(3):788–797.

[Shapley and Perry, 1986] Shapley, R. & Perry, V. H. (1986). Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neurosciences*, 9:229–235.

[Sherman and Guillery, 2006] Sherman, S. M. & Guillery, R. W. (2006). *Exploring the thalamus and its role in cortical function*. MIT Press.

[Shou et al., 2000] Shou, T., Wang, W., & Yu, H. (2000). Orientation biased extended surround of the receptive field of cat retinal ganglion cells. *Neuroscience*, 98(2):207–212.

[Silveira and Perry, 1991] Silveira, L. & Perry, V. (1991). The topography of magnocellular projecting ganglion cells (m-ganglion cells) in the primate retina. *Neuroscience*, 40(1):217–237.

[Solomon et al., 2002] Solomon, S. G., White, A. J., & Martin, P. R. (2002). Extraclassical receptive field properties of parvocellular, magnocellular, and koniocellular cells in the primate lateral geniculate nucleus. *Journal of Neuroscience*, 22(1):338–349.

[Spehar et al., 2003] Spehar, B., Clifford, C. W., Newell, B. R., & Taylor, R. P. (2003). Universal aesthetic of fractals. *Computers & Graphics, 27*(5), 813-820.

[Spehar et al., 2015] Spehar, B., Wong, S., van de Klundert, S., Lui, J., Clifford, C. W. G., & Taylor, R. (2015). Beauty and the beholder: the role of visual sensitivity in visual preference. *Frontiers in human neuroscience, 9*, 514.

[Spillmann, 1994] Spillmann, L. (1994). The Hermann grid illusion: A tool for studying human perspective field organization. *Perception, 23*, 691–708.

[Todorovic, 1997] Todorovic, D. (1997). Lightness and junctions. *Perception*, 26, 379–395.

[Tomasi and Manduchi, 1998] Tomasi, C., & Manduchi, R. (1998, January). Bilateral filtering for gray and color images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)* (pp. 839-846). IEEE.

[Ungerleider and Mishkin, 1982] Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In: *The Analysis of Visual Behavior*, ed. D. J. Ingle, R. J. W. Mansfield, M. S. Goodale, pp. 54-86. Cambridge, Mass: MIT Press.

[Van De Ville and Unser, 2008] Van De Ville, D., & Unser, M. (2008). Complex wavelet bases, steerability, and the Marr-like pyramid. *IEEE Transactions on Image Processing*, 17(11):2063-2080.

[White, 1979] White, M. (1979). A new effect of pattern on perceived lightness. *Perception*, 8(4):413–416.

[White, 1981] White, M. (1981). The effect of the nature of the surround on the perceived lightness of grey bars within square-wave test gratings. *Perception*, 10(2):215–230.

[Wind, 1899] Wind, C.H.: Zur demonstration einer von E. Mach entdeckten optischen Täuschung. *Physik. Zeit. 1*, 112–113 (1899)

[Xu et al., 2001] Xu, X., Ichida, J. M., Allison, J. D., Boyd, J. D., Bonds, A., & Casagrande, V. A. (2001). A comparison of koniocellular, magnocellular and parvocellular receptive field properties in the lateral geniculate nucleus of the owl monkey (aotus trivirgatus). *The Journal of Physiology*, 531(1):203–218.

[Young, 1987] Young, R. A. (1987). The gaussian derivative model for spatial vision: I. retinal mechanisms. *Spatial Vision*, 2(4):273–293.

Relevant Publications

•Journals:

—Bakshi, A. & Ghosh, K. (2012). Some insights into why the perception of mach bands is strong for luminance ramps and weak or vanishing for luminance steps. *Perception*, 41(11):1403–1408.

—Bhattacharjee, D., Bakshi, A., & Ghosh, K. (2015). Comparison between an HVS inspired linear filter and the bilateral filter in performing "vision at a glance" through smoothing with edge preservation. *International Journal of Image and Graphics*, 15(04):15500-15.

—Bakshi, A., Roy, S., Mallick, A., & Ghosh, K. (2016). Limitations of the oriented difference of gaussian filter in special cases of brightness perception illusions. *Perception*, 45(3):328-336.

—Bakshi, A., Roy, S., Mallick, A., & Ghosh, K. (2017) Experimental Observations of the Visibility Threshold of Illusory effects in Hermann Grid, Sinusoidal and Square Gratings and their Possible Implications. *Int J Open Access Ophthal* 2(1):1-8. DOI: 10.15226/2474-9249/2/1/00120

—Bakshi, A., & Ghosh, K. (2018). A parsimonious model of brightness induction. *Biological Cybernetics*, 112(3):237-251.

—Bakshi, A., & Ghosh, K. (2020). Tiny Squares at the Hermann Grid Corners Can Completely Remove the Illusion. *Perception*, *49*(2), 232-239.

—Bakshi, A., Roy, S., Mallick, A., & Ghosh, K. (2021). A discrete magno–parvo additive model in early vision for explaining brightness perception in varying contrastive contexts. *Biological Cybernetics*, (published online: 1-17). https://doi.org/10.1007/s00422-021-00896-4

•Conferences:

—*Perception and Machine Intelligence*, *First Indo-Japan Conference, PerMIn 2012*, 12-13 Jan. 2012, Kolkata, Bakshi A, & Ghosh K, "Scaling properties of Mach bands and perceptual models", In: Kundu M.K., Mitra S., Mazumdar D., Pal S.K. (eds) Perception and Machine Intelligence. PerMIn 2012. Lecture Notes in Computer Science, vol 7143, pp. 66-74. Springer, Berlin, Heidelberg

—*International Advance Computing Conference (IACC), 2013, IEEE 3rd International* , 22-23 Feb. 2013, Ghaziabad, Dasgupta, A.; Bakshi, A.; & Ghosh, K., "Lateral inhibition based holistic approach to adaptive image enhancement," pp.1143-1148

—*5th World Conference on Paraconsistency, Kolkata, 2014,* Bakshi A., & Ghosh K. (2015) Perceiving and Modeling Brightness Contradictions Through the Study of Brightness Illusions. In: Beziau JY., Chakraborty M., Dutta S. (eds) New Directions in Paraconsistent Logic. Springer Proceedings in Mathematics & Statistics, vol 152, pp. 447-463. Springer, New Delhi

—*37th European Conference on Visual Perception (ECVP), 2014*, Kuntal Ghosh, Ashish Bakshi, Arijit Mallick & Sourya Roy, "Experimental Studies on the Sinusoidal Grating, Square Grating and Hermann Grid Illusions for Visibility Thresholds", August 24-28, 2014, Belgrade, Serbia, Poster presentation & Abstract Publication in *Perception* Vol. 43, pp. 131-132. https://doi.org/10.1177/03010066140430S101

—*38th European Conference on Visual Perception (ECVP)*, *2015*, Kuntal Ghosh, Arijit Mallick, Sourya Roy & Ashish Bakshi, "Limitations of the ODOG filter in special cases of brightness

perception illusions", Aug 23-27, 2015, Liverpool, United Kingdom, Poster presentation & Abstract Publication in *Perception* Vol. 44, pp. 252-252.
https://doi.org/10.1177/0301006615598674

—*39th European Conference on Visual Perception (ECVP)*, *2016*, Kuntal Ghosh, Ashish Bakshi, Sourya Roy & Arijit Mallick, "Attention as a new parameter in modeling brightness induction", Aug 28-Sep 01, 2016, Barcelona, Spain, Poster presentation & Abstract Publication in *Perception* Vol. 45, pp. 133-133.
https://doi.org/10.1177/0301006616671273

- **Published as book chapter:**

Ashish Bakshi, Kuntal Ghosh, "A Neural Model of Attention and Feedback for Computing Perceived Brightness in Vision." In: Samui, P., Roy, S. S., & Balas, V. E. (Eds.). (2017). *Handbook of Neural Computation*, pp. 487-513. Academic Press.