
Contours and Contrast

Kaleigh Smith

**Max-Planck-Institut für Informatik
Saarbrücken, Germany**

A thesis for obtaining the title of Doctor of Engineering
of the Faculties of Natural Sciences and Technology of Saarland University.
December 11th, 2008
Saarbrücken, Germany

Dekan — Dean

Prof. Dr. Joachim Weickert, Universität des Saarlandes, Saarbrücken, Germany

Betreuender Hochschullehrer — Supervisor

Dr.-Ing. Habil. Karol Myszkowski, MPI für Informatik, Saarbrücken, Germany

Gutachter — Reviewers

Prof. Dr. Hans-Peter Seidel, MPI für Informatik, Saarbrücken, Germany

Dr.-Ing. Habil. Karol Myszkowski, MPI für Informatik, Saarbrücken, Germany

Prof. Dr. Joëlle Thollot, Universités Grenoble, Grenoble, France

Datum des Kolloquiums — Date of Defense

11. December 2008

Kolloquium — Members of the Examination Board

Prof. Dr. Joachim Weickert, Universität des Saarlandes - Chair

Dr.-Ing. Habil. Karol Myszkowski, MPI für Informatik - Examiner

Prof. Dr. Hans-Peter Seidel, MPI für Informatik - Examiner

Prof. Dr. Habil. Joëlle Thollot, Universités Grenoble - Examiner

Dr.-Ing. Thorsten Grosch, MPI für Informatik - Reporter

Kaleigh Smith

Max-Planck-Institut für Informatik

Stuhlsatzenhausweg 85

66123 Saarbrücken, Germany

kaleigh.smith@gmail.com

Abstract

Contrast in photographic and computer-generated imagery communicates colour and lightness differences that would be perceived when viewing the represented scene. Due to depiction constraints, the amount of displayable contrast is limited, reducing the image's ability to accurately represent the scene. A local contrast enhancement technique called unsharp masking can overcome these constraints by adding high-frequency contours to an image that increase its apparent contrast. In three novel algorithms inspired by unsharp masking, specialized local contrast enhancements are shown to overcome constraints of a limited dynamic range, overcome an achromatic palette, and to improve the rendering of 3D shapes and scenes. The *Beyond Tone Mapping* approach restores original HDR contrast to its tone mapped LDR counterpart by adding high-frequency colour contours to the LDR image while preserving its luminance. *Apparent Greyscale* is a multi-scale two-step technique that first converts colour images and video to greyscale according to their chromatic lightness, then restores diminished colour contrast with high-frequency luminance contours. Finally, *3D Unsharp Masking* performs scene coherent enhancement by introducing 3D high-frequency luminance contours to emphasize the details, shapes, tonal range and spatial organization of a 3D scene within the rendering pipeline. As a perceptual justification, it is argued that a local contrast enhancement made with unsharp masking is related to the Cornsweet illusion, and that this may explain its effect on apparent contrast.

Kurzfassung

Inspiziert von Unsharp Masking, werden drei Algorithmen präsentiert, welche zeigen, dass spezialisierte lokale Kontraststeigerungen die Einschränkungen des dynamischen Wertebereichs, sowie die einer achromatischen Farbpalette überwinden, und das Rendern von 3D Objekten und Szenen verbessern. Der "Beyond Tonemapping" Ansatz stellt den originalen HDR Kontrast in seinem farbadaptierten LDR Gegenstück wieder her, indem er hoch-frequente Farbkonturen zu dem LDR Bild hinzufügt und dabei die Luminanz erhält. "Apparent Greyscale" ist eine Multiskalen Technik, die Farbbilder und Videos in Graustufen konvertiert, indem der Farbkontrast mittels hoch-frequenter Luminanz-Konturen wieder hergestellt wird. Zu letzte führt "3D Unsharp Masking" eine szenenkohärente Bildverbesserung durch, indem hoch-frequente 3D Luminanz-Konturen eingeführt werden, um die Details, Objekte, Farbbereiche, und die räumliche Organisation in der Rendering Pipeline von 3D Szenen hervor zu heben. Als wahrnehmungsbasierte Rechtfertigung wird dargelegt, dass Kontraststeigerungen mittels Unsharp Masking mit der Cornsweet Illusion verwand sind und erklären könnte, warum das Hinzufügen hoch-frequenter Konturen zu einem Bild seinen anscheinenden Kontrast erhöht.

Summary

Contrast in photographic and computer-generated imagery communicates colour and lightness differences that would be perceived when viewing the represented scene. Due to depiction constraints, the amount of displayable contrast is limited, reducing the image's ability to accurately represent the scene. Local contrast enhancement can overcome these constraints to produce improved imagery with higher information content, resulting in more efficient depictions. One such technique, called unsharp masking, influences the perception of contrast by adding high-frequency contours to an image.

In three novel algorithms inspired by unsharp masking, local contrast enhancement is shown to overcome a limited dynamic range, overcome an achromatic palette, and to improve the rendering of 3D shapes and scenes. The *Beyond Tone Mapping* approach restores original HDR contrast to its tone mapped LDR counterpart by adding high-frequency colour contours to the LDR image but preserving its luminance. *Apparent Greyscale* is a multi-scale two-step technique to convert colour imagery to greyscale that restores diminished colour contrast with high-frequency luminance contours. Finally, *3D Unsharp Masking* performs scene coherent enhancement in a technique that introduces high-frequency 3D luminance contours to emphasize the details, shapes, tonal range and spatial organization of a 3D scene within the rendering pipeline.

The algorithms are justified by perceptual explanations of actual and apparent contrast. In particular, it is argued that local contrast enhancements made with unsharp masking can be explained by the Cornsweet illusion. This illusion creates a perceived brightening or darkening of regions that are adjacent to a specially shaped high-frequency contour, much like those that are introduced through unsharp masking. The same perceptual mechanism may explain why contours added to an image increase its apparent contrast.

Restoring Apparent Contrast to Tone Mapped Images

High Dynamic Range (HDR) images capture the full range of luminance present in real world scenes, and unlike Low Dynamic Range (LDR) images, can simultaneously contain detailed information in the deepest of shadows and the brightest of light sources. For display or aesthetic purposes, it is often necessary to perform tone mapping, which creates LDR depictions of HDR images at the cost of contrast information loss. The *Beyond Tone Mapping* approach enhances an LDR depiction with colour adjustments that add base layer and detail layer contrast without modifying luminance. Its adaptation of traditional unsharp masking calculates the local HDR contrast lost during tone mapping, then restores it to the LDR chromatic channels.

Greyscale Conversion of Images and Video

Although colour printing has become common practice, artists and publishers continue to employ greyscale. The format is more dependable and often more evocative. One of the most basic tools in digital image editing software is the greyscale converter, which takes a colour image and produces a colourless version. The depiction challenge is to ensure that the original's chromatic contrasts are communicated even when no colour is present. *Apparent Greyscale* is a quick and simple method for converting complex images and video to perceptually accurate greyscale versions. It is a two-step approach to globally assign grey values and determine colour ordering, that then locally

enhances the greyscale to reproduce the original contrast. The global mapping is image independent and incorporates the Helmholtz-Kohlrausch colour appearance effect for predicting differences between isoluminant colours. The multiscale local contrast enhancement reintroduces lost discontinuities only in regions that insufficiently represent original chromatic contrast. All operations are restricted so that they preserve the overall image appearance, lightness range and differences, colour ordering, and spatial details, resulting in perceptually accurate achromatic reproductions of the colour original.

Enhanced 3D Rendering

In 3D rendering, the virtual camera settings are fixed, meaning that not all relevant scene information may be visible. The *3D Unsharp Masking* approach enhances local scene contrast by unsharp masking over arbitrary surfaces under any form of illumination. The adaptation of the well-known 2D unsharp masking technique to 3D interactive scenarios is designed to aid viewers in tasks like understanding complex or detailed geometric models, medical visualization and navigation in virtual environments. Its holistic approach enhances the depiction of various visual cues, including gradients from surface shading, surface reflectance, shadows, and highlights, to ease estimation of viewpoint, lighting conditions, shapes of objects and their world-space organization. The operation runs at real-time rates on a GPU and the effect is easily controlled interactively within the rendering pipeline. It is validated by psychophysical experiments showing that the enhanced images are perceived as having better contrast and are preferred over unenhanced originals.

Zusammenfassung

Kontrast in Photographie und computergenerierten Bildern vermittelt Farb- und Helligkeitsunterschiede, die beim Ansehen einer Szene wahrgenommen werden würden. Auf Grund von Beschränkungen bezüglich der Displaykapazitäten ist der darstellbare Kontrast limitiert, wodurch die Fähigkeit eines Bildes die Szene akkurat zu repräsentieren reduziert wird. Aus diesem Grund sind effektive Methoden zur Kontrastdarstellung sowohl für Bildverarbeitung als auch für das Nachbearbeiten von Bildern erforderlich.

Glücklicherweise können lokale Kontraststeigerungen diese Beschränkung überwinden, um Bilder mit erhöhtem Informationsgehalt zu produzieren, was in einer effizienteren Darstellung resultiert. Eine solche Methode, Unsharp Masking, beeinflusst die Wahrnehmung von Kontrast, indem hoch-frequente Konturen zu einem Bild hinzugefügt werden.

Inspiriert von Unsharp Masking, werden drei Algorithmen präsentiert, die zeigen, dass lokale Kontrastverbesserungen die Einschränkungen des dynamischen Wertebereichs, sowie die einer achromatischen Farbpalette überwinden, und das Rendern von 3D Objekten und Szenen verbessern.

Der “Beyond Tonemapping” Ansatz stellt den originalen HDR Kontrast in seinem farbadaptierten LDR Gegenstück wieder her, indem er hoch-frequente Farbkonturen zu dem LDR Bild hinzufügt und dabei die Luminanz erhält. “Apparent Greyscale” ist eine Multiskalen Technik, um Farbbilder und Videos in Graustufen zu konvertieren, indem der Farbkontrast mittels hoch-frequenter Luminanz-Konturen wieder hergestellt wird. Zu letzte führt “3D Unsharp Masking” eine szenenkohärente Bildverbesserung durch, indem hoch-frequente 3D Luminanz-Konturen eingeführt werden, um die Details, Objekte, Farbbereiche, und die räumliche Organisation in der Rendering Pipeline von 3D Szenen hervor zu heben.

Diese Algorithmen werden durch wahrnehmungsbasierte Erklärungen bezüglich des tatsächlichen und anscheinenden Kontrasts gerechtfertigt. Im Besonderen wird argumentiert, dass lokale Kontraststeigerungen mittels Unsharp Masking durch die Cornsweet Illusion erklärt werden kann.

Diese Illusion erzeugt ein erkennbares Aufhellen und Verdunkeln in Regionen, die entlang hoch-frequenter Konturen verlaufen, ähnlich derer die durch Unsharp Masking generiert werden, und der gleiche erkennbare Effekt könnte erklären, warum zu einen Bild hinzugefügte Konturen den anscheinenden Kontrast erhöhen.

Acknowledgements

First, I'd like to thank my co-authors Pierre-Edouard Landes and Tobias Ritschel for their readiness to try new ideas, their love for designing software and for indulging my insistence that images can always look better. To Thorsten Grosch for his careful reading and editing. To Matthias Ihrke for running a valuable user study on the Cornsweet effect in 3D.

I owe much to the many professors and researchers who I've had the fortune to work with. To Allison Klein, who introduced me to computer graphics and put me through my first SIGGRAPH challenge. To Victor Ostromoukhov, who encouraged me to continue my doctoral work and facilitated my move to MPI. To George Dretakis, whose guidance safely steered my research to publication and whose faith in my abilities was a much needed encouragement. To Joëlle Thollot, you are a remarkable woman and supervisor - you've selflessly nurtured my research, have accepted my failings, and shown by example what it is to be a scientist, artist, mentor and mediator. To Karol Myszkowski, my supervisor, for introducing me to the Cornsweet illusion and colour theory, and especially for his patience and sound advice. I am forever grateful for the freedom you have given me in my research - for never saying "no".

To Hans-Peter Seidel, MPI, IMPRS and all the administrative people I have irritated in Saarbrücken with mis-filled forms and special requests. I appreciate the travel opportunities, good working environment and leave for both my research visit to France and internship at Google. To the ARTIS/AERIS group and everyone at INRIA Grenoble, you are a lovely team: supportive, encouraging and fun. To everyone who edited something I've written, and to the reviewers of various computer graphics conferences and journals, your comments and advice are invaluable.

To Margaret Livingstone, whose book on vision and art inspired me to look at graphics with an artistic eye. To Floyd Raliff, whose pioneering work on lightness and colour phenomena and their connections to artistic techniques was, and still is, an inspiration. His writing is admirable, and he authored the only paper that, over three long years of paper reading, made me laugh.

To everyone on my email list who participated in many user studies, my apologies that none of that research is presented here. To Saschka Unseld for the mice and encouragement. To my Saarbrücken supporters, for yoga, dinners, travel and lifetime friendships. To Tom Annen for treats, laughs, German translation, and for being an amazing colleague and cherished friend. To my parents and sister, for their continued love and support of a perpetual adult student.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 1.1 | Contrast, Contours and Cornsweet | 1 |
| 1.2 | Depiction Despite Constraints | 5 |
| 1.3 | Artistic Motivation | 8 |
| 1.4 | Novel Contributions | 9 |
| 1.5 | Thesis Organization | 11 |
| 2 | Contrast Perception and Enhancement | 13 |
| 2.1 | Seeing Contrast | 14 |
| 2.1.1 | Lightness constancy | 14 |
| 2.1.2 | Spatial Vision | 15 |
| 2.1.3 | Spatial Frequency and Contrast Sensitivity | 17 |
| 2.1.4 | Apparent Contrast | 18 |
| 2.2 | Measuring Contrast | 19 |
| 2.2.1 | Local Contrast | 20 |
| 2.2.2 | Local Band-limited Contrast | 22 |
| 2.3 | Chromatic Contrast | 24 |
| 2.3.1 | Colour Model Background | 24 |
| 2.3.2 | CIE $L^*a^*b^*$ and $L^*u^*v^*$ | 25 |
| 2.3.3 | The Helmholtz-Kohlrausch Effect | 27 |
| 2.3.4 | Survey of Chromatic Lightness Predictors | 28 |
| 2.4 | Contrast Enhancement | 29 |
| 2.4.1 | Global Enhancement | 29 |
| 2.4.2 | Local Enhancement | 30 |
| 2.4.3 | Edge Sharpening Filters | 31 |
| 2.4.4 | Unsharp Masking | 32 |
| 2.5 | Preferred Contrast and Colour Saturation | 34 |
| 3 | Contour Brightness Illusions | 37 |
| 3.1 | Contours and Apparent Contrast | 38 |
| 3.2 | The Cornsweet Illusion | 39 |
| 3.2.1 | Perceived Strength of the Cornsweet Illusion | 40 |
| 3.2.2 | The Chromatic Cornsweet Illusion | 42 |
| 3.2.3 | The 3D Cornsweet Illusion | 43 |
| 3.3 | Perceptual Explanations | 45 |
| 3.4 | Unsharp Masking and the Cornsweet Illusion | 46 |

| | | |
|----------|---|------------|
| 4 | Restoring Apparent Contrast To Tone Mapped Images | 49 |
| 4.1 | In Art: Overcoming a Limited Dynamic Range | 50 |
| 4.2 | Tone Mapping Background | 51 |
| 4.3 | Beyond Tone Mapping | 52 |
| 4.3.1 | Colour Contrast and Chromatic Scaling | 53 |
| 4.3.2 | Distortion Method | 55 |
| 4.3.3 | Unsharp Method | 57 |
| 4.3.4 | Resulting Images | 58 |
| 4.4 | Summary and Discussion | 60 |
| 5 | Greyscale Conversion of Images and Video | 67 |
| 5.1 | In Art: Contours Overcome An Achromatic Palette | 67 |
| 5.2 | Survey of Greyscale Conversion Methods | 68 |
| 5.3 | Apparent Greyscale | 71 |
| 5.3.1 | Global Mapping to Apparent Lightness | 71 |
| 5.3.2 | Local Chromatic Contrast Enhancement | 74 |
| 5.4 | Applications of Apparent Greyscale | 80 |
| 5.4.1 | Isoluminant Images | 80 |
| 5.4.2 | Image Sets and Video | 80 |
| 5.4.3 | Preserving Spatial Details and Salient Regions | 82 |
| 5.5 | Implementation and Limitations | 87 |
| 5.6 | Summary and Discussion | 89 |
| 6 | Enhanced 3D Rendering | 91 |
| 6.1 | In Art: Contours Emphasize Shape and Scene | 91 |
| 6.2 | Recent Approaches to Enhance 3D Graphics | 93 |
| 6.3 | 3D Unsharp Masking | 95 |
| 6.3.1 | Extension from 2D to 3D Unsharp | 96 |
| 6.3.2 | Preserving Chromatic Saturation Levels | 96 |
| 6.3.3 | Adaptive Gain | 97 |
| 6.3.4 | Contrast Signal Calculation | 97 |
| 6.3.5 | Comparison of 2D and 3D Unsharp Masking | 98 |
| 6.3.6 | Implementation | 100 |
| 6.3.7 | Limitations | 100 |
| 6.4 | Applications of 3D Unsharp Masking | 101 |
| 6.5 | Perceived Effect of 3D Unsharp Masking | 108 |
| 6.5.1 | Perceived Contrast Enhancement | 108 |
| 6.5.2 | 3D Cornsweet Illusion | 110 |
| 6.6 | Summary and Discussion | 110 |
| 7 | Conclusion | 113 |
| 7.1 | Summary of Contributions | 114 |
| 7.2 | The Future of Contours in CG | 115 |
| 7.3 | Concluding Remarks | 115 |

Preface

This thesis documents an investigation into the effect of contours on contrast perception and its application to computer depiction. Justified by findings in visual perception and demonstrated through novel graphics techniques, the conclusion is that local contrast alteration is a powerful Computer Graphics tool that is useful in a variety of situations. The tone of this thesis is unlike that of the papers from which it arose. In addition to championing individual applications, there is an emphasis on contrast perception and on brightness phenomena created by contours. The intended audience is anyone interested in applying perception to computer graphics, and researchers working on tone mapping, greyscale conversion or enhanced 3D rendering.

Chapter 1

Introduction

“When a dog bites a man, that is not news... But if a man bites a dog, that is news”[...] The relation between contour and contrast is similar. It is not news, but old and common knowledge that it is the contrast, or difference, in the brightness or color of two adjacent areas that causes the appearance of a contours between them. [...] We shall then consider the “news”: how contour itself can affect the contrast of adjacent areas...

– Floyd Ratliff (1919-1999)

1.1 Contrast, Contours and Cornsweet

Whether indoors, outside, looking at a photograph or watching a video, it is differences in brightness or colourfulness that allow us to make sense of the visual stimulus with which we are presented. This difference, called *contrast*, forms the core of visual processing. It allows a viewer to differentiate surfaces, discern object boundaries, interpret depth relationships and ultimately, to understand what he or she is seeing.

Humans continuously adapt their eyes to their surroundings so that there is sufficient incoming contrast to produce a response from the visual system. Named *brightness adaptation*, this means that in shadow, the eye becomes more sensitive to light so as to discern texture details and geometry discontinuities, whereas in brightly lit areas, the eye is less sensitive so that it is able to perceive bright shading gradients and details around highlights. Despite these adjustments, contrast is perceived consistently over an entire scene. In natural cases, humans are capable of discounting luminance differences caused by illumination to properly infer lighting changes due to surface geometry or material. Thus, seeing is not a simple task of measuring actual luminance, but is a translation from incoming light to a coherent impression of brightness and contrasts between regions.

Images and video, whether, painted, photographed or rendered, all strive to depict a scene’s contrast. They act as stand-ins for the actual scene, attempting to create a similar impression of its features and contrasts. However, due to the display medium, tone and colour range, camera aperture and exposure and a host of other reasons, their contrast is diminished. For instance, when creating a displayable image, the full tonal range of the original scene must be compressed, which can lead to losses in global and

local contrast. When converting a colour image to greyscale, some chromatic contrasts become faint or invisible. And when rendering a scene with fixed lighting and a set aperture, not all shapes and materials are visible; some disappear into dark regions, other become washed out or clamped to the highest value. Thus contrast depiction is a great challenge to image processing and rendering.

Contrast Depiction The visual communication of all the important contrasts making up a real or synthetic scene. The challenge is to create an image that overcomes the constraints imposed by the depiction medium.

Contrast in an image can be defined in various ways, but in principle it is a difference in luminance or colour between two locations. Global contrast, or overall contrast, is generally defined as the difference between the brightest and darkest regions in the image. It depicts the luminance range of the original scene and helps to identify light sources and distinguish the figures or foreground from the background. Considered locally, contrast is the difference between adjacent regions and its magnitude is usually measured by comparing a small region to its neighbouring regions. Local contrast aids in the perception of details, textures and spatial arrangement.

The transition between two contrasting adjacent regions is called a *contour*. The term contour is used because it captures the notion of a transition that may extend over an area, whereas the term *edge* usually means a collection of precise maximum or minima that occur at single points. Contours delineate the differing areas of an image and are capable of communicating the scene from which they arise. For instance, line drawings depicting contours are our earliest and most efficient means of reproducing visual information. In fact, they are specially processed to the point that humans are sensitive to illusory contours that are not depicted, but implied by surrounding shapes [Palmer, 1999].

In addition to its difference magnitude, a contour may be very sharp, or more gradual. An abrupt contour makes for a sharper contrast, and a wider contour makes for a smoother change. Thus, its prominence depends both on the difference between its two adjoining regions and on the spatial extent of the transition. What is remarkable, and is the foundation upon which this research stands, is that contours themselves have an effect on apparent contrast; they are not only products of contrast, but produce it of their own accord.

Such an occurrence is due to local context; the percept of a bright stimulus appears brighter in comparison to something dark. This *apparent contrast* is a perceived difference, as opposed to a physically measurable difference.

Apparent Contrast A perceptual measure of contrast defined as a difference or rate of change in brightness, where brightness is the perceived amount of light emitted from an area.

In studying contrast perception to learn about visual processing, researchers have observed various brightness phenomena attributed to contours. Particular interest has been paid to the Craik-Cornsweet-O'Brien family of illusions, in which specifically shaped contours create the illusion of brightness differences. These illusions occur when a visual stimulus garners a visual response that differs significantly from its luminance signal. For example, looking at Figure 1.1(a), a human is tricked into perceiving a step edge of actual contrast, like that shown in Figure 1.1(b). This so-called *Cornsweet illusion* is caused by a *Cornsweet contour*, whose luminance profile consists of sharp opposing peaks that quickly return to the same luminance. The contour

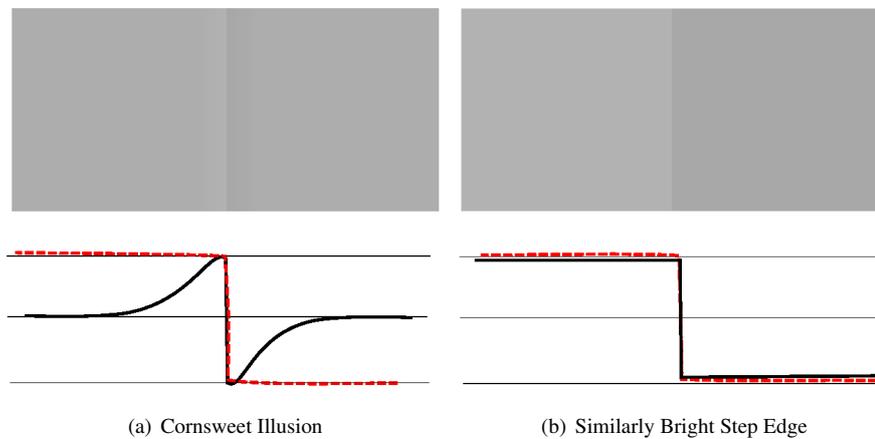


Figure 1.1: The Cornsweet Illusion (a), and a step edge of similar luminance and contrast (b). Luminance profiles are in solid black and the approximate perceived brightness is in dashed red. Notice the large increase in apparent contrast (range of dashed red line) over the objectively measured contrast (range of black line) of the Cornsweet illusion (a).

is perceived as an actual brightness difference between adjacent regions, creating the illusion of local contrast when in fact there is none [Cornsweet, 1970].

Cornsweet Contour A contour whose luminance profile of sharp opposing peaks gradually returns to the same luminance, or to luminances of lesser contrast.

From a Computer Graphics perspective, this illusion presents an opportunity for controlled and localized image adjustment that can solve contrast-related problems and to potentially reverse the effects of contrast loss. Can a Cornsweet contour be used to make images that are more effective and informative? Consider the edge stimulus shown in Figure 1.2(a) and suppose the light and dark regions are not as light or dark as they should be. Overlaying a Cornsweet contour on the existing edge adjusts the luminance profile by modifying the values near the edge, but leaving the rest unchanged, as illustrated in Figure 1.2(b). The perceived brightness of both sides has changed, adding enough apparent contrast to make the image appear the same as the higher contrast step edge in Figure 1.2(c). This simple example illustrates the first principle of this book:

First Principle Adding a Cornsweet contour can increase apparent contrast beyond the physical contrast in complex images.

Consider the luminance profile of a Cornsweet contour whose values rise or dip near its high-frequency edge, then fall back to match the low-frequency values away from the edge. This type of profile could be created when the amplitude of existing high-frequencies increase, or when high-frequencies from a step edge are introduced. In image enhancement, scaling high-frequencies that fall within a certain range makes discontinuities in that range appear more pronounced, and is used to sharpen edges and increase detail visibility. One such well-known technique is called *unsharp masking*, in which an original signal is enhanced by the addition of a *contrast signal* that is scaled by a *gain factor* usually termed λ [Badamchizadeh and Aghagolzadeh, 2004].

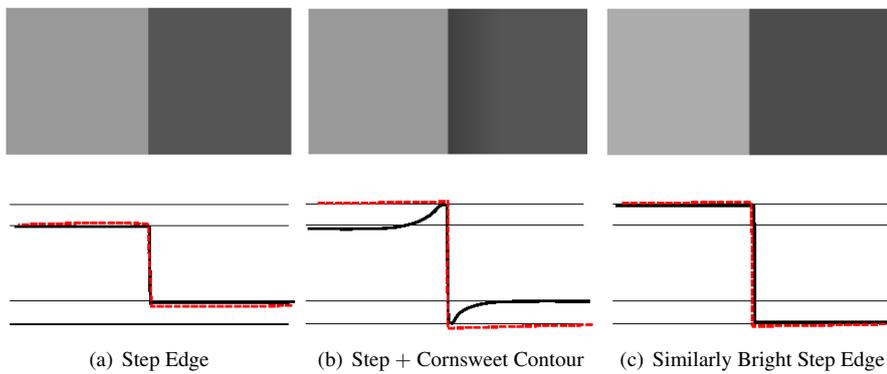


Figure 1.2: An illustration of the first principle: adding a Cornsweet contour to an existing step edge changes its luminance profile slightly (black line) and its brightness profile (red dashed line) more drastically.

The first observation to be made is that Cornsweet contours can be created using a type of unsharp masking, as illustrated in Figure 1.3. The second observation is that the Cornsweet illusion may be a perceptual explanation for the contrast enhancing power of the technique. These observations form the second principle of this book:

Second Principle Unsharp masking is capable of introducing Cornsweet contours, and the perceptual effect of unsharp masking can be explained by the Cornsweet illusion.

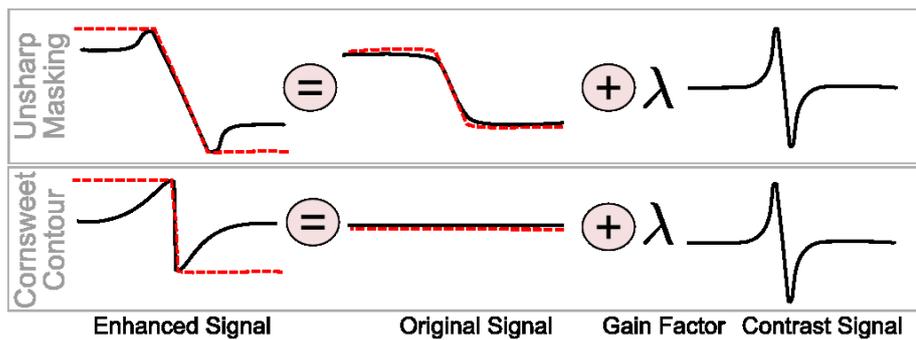


Figure 1.3: Diagram of the relationship between unsharp masking of an edge and the creation of a Cornsweet contour that causes a brightness illusion. Both introduce scaled high-frequencies to create an enhanced version of the original, which in turn affects both the real contrast (range of solid black lines) and apparent contrast (range of red dashed lines).

1.2 Depiction Despite Constraints

Let us return to the main problem addressed here: the depiction of contrast despite constraints. Focus is placed on three topics in graphics, each embodying a different depiction challenge, but all sharing the common goal of representing information despite conditions that restrict the full original information from being depicted.

The first topic, *Tone Mapping*, considers how images with a restricted range of dark to light values can depict the real world’s seemingly unlimited range. *Greyscale Conversion* is a topic that investigates how black and white imagery can be created to reproduce the appearance of the full colour originals. The third topic is *3D Rendering*, which generates rendered imagery under static lighting and a fixed aperture camera with enough contrast to sufficiently depict the shape, spatial arrangements and surface details of a 3D scene.

There are diverse ways to create images and video that solve depiction problems within each of these three areas. Some involve optimizations and gradient domain processing, others simulate real-world solutions, and others still modify the data heuristically to improve the depiction. Contrast-related manipulations similar to unsharp masking, where contrast is measured at a chosen scale and increased, are popular in image enhancement and have been proven successful as making local and adaptive changes [Peli, 2000].

As summarized in Table 1.2, the purpose of this thesis is to discuss each of these three topics and to present novel contrast-related solutions based on unsharp masking. These novel approaches create superior imagery by adding specific lost or under-represented contrast to images and video.

| Tone Mapping | Greyscale Conversion | 3D Rendering |
|----------------------------|-----------------------------|---------------------------|
| <i>Beyond Tone Mapping</i> | <i>Apparent Greyscale</i> | <i>3D Unsharp Masking</i> |
| Restore lost contrast | Preserve chromatic contrast | Enhance scene contrast |
| HDR/LDR image pair | Colour image/video | 3D scene |
| ↓ | ↓ | ↓ |
| LDR | Greyscale image/video | Rendered image/video |

Table 1.1: Overview of the topics addressed in this thesis and the purpose, input and output of the novel Cornsweet contour approaches. The names of the novel approaches for depicting contrast despite constraints are italicized.

Tone Mapping

High Dynamic Range (HDR) images capture the full range of luminance present in real world scenes, and unlike Low Dynamic Range (LDR) images, can simultaneously contain detailed information in the deepest of shadows and the brightest of light sources. An HDR photograph can be captured using a special camera or by combining a set of photographs with varying exposure settings, like those shown in Figure 1.4(a). For display or aesthetic purposes, it is often necessary to perform tone mapping, which creates LDR depictions of HDR images at the cost of contrast compression, which results in information loss. This compression is done using a tone mapping operator and produces vivid imagery, such as the LDR image shown in Figure 1.4(b), which resulted from gradient domain compression [Fattal *et al.*, 2002].

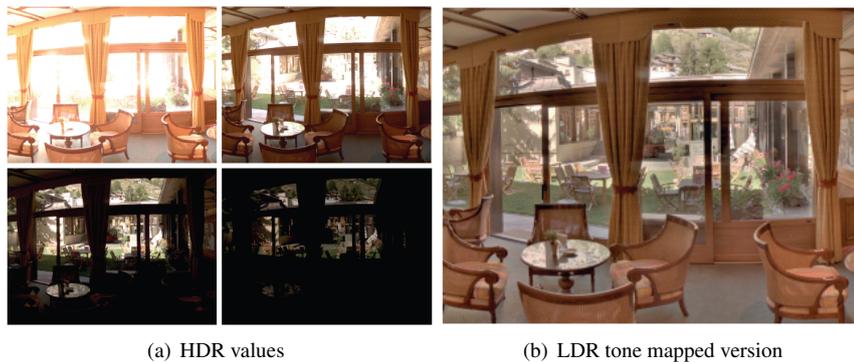


Figure 1.4: View of different value ranges within a single HDR image (left, 4 different exposures); An LDR version resulting from the gradient domain compression tone mapper [Fattal *et al.*, 2002] which combines the HDR range of values into one single displayable luminosity range (right).

The two *Beyond Tone Mapping* algorithms presented here enhance an LDR depiction with perceptually driven colour adjustments to restore the original HDR contrast information. Artistic and photographic colour techniques are identified, from which adjustments that create contrast with colour are derived. The enhanced LDR image is an improved depiction of the original HDR image with restored contrast information. The first algorithm, named the *distortion method*, takes distortion measurements between the HDR and LDR pair, then performs enhancement of LDR depictions to either emphasize details or aid in figure/ground separation, Figure 1.5(b). The second algorithm, named the *unsharp method*, is a more automatic version of the first, and proceeds by directly measuring contrast loss at a base and detail scale, Figure 1.5(c). This loss then drives chromatic scaling which increases apparent contrast.



Figure 1.5: The Café image resulting from *photoreceptor* tone mapping [Reinhard and Devlin, 2005] (a), and with restored contrast from the two *Beyond Tone Mapping* methods (b & c). Notice that the flowers, chairs and umbrellas are more visible and the distant landscape contains more details and depth.

Greyscale Conversion

Although colour printing has become common practice, artists and publishers continue to employ greyscale. The format is more dependable and often more evocative. One of the most basic tools in digital image editing software is the greyscale converter, which

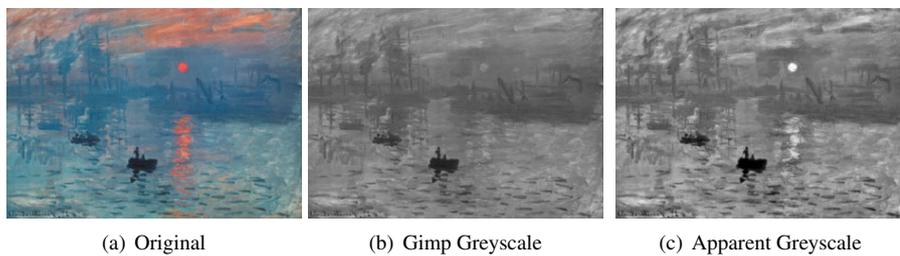


Figure 1.6: A photo of Claude Monet’s “Impression Sunrise” painting (a). An example of a standard conversion from an original colour photograph to a black and white version created by *The Gimp*’s greyscale mode (b). The novel *Apparent Greyscale* approach produces (c) by maintaining image appearance and preserving chromatic contrast.

takes a colour image like Figure 1.6(a), and produces a colourless version like Figure 1.6(b). The depiction challenge is to ensure that the original’s chromatic contrasts are communicated even when no colour is present.

Apparent Greyscale is a quick and simple method for converting complex images and video to perceptually accurate greyscale versions. Its two-step approach first globally assigns grey values and determines colour ordering, then second, locally enhances the greyscale to reproduce the original contrast. The global mapping is image independent and incorporates the Helmholtz-Kohlrausch colour appearance effect for predicting differences in perceived brightness between isoluminant colours. The multiscale local contrast enhancement reintroduces lost discontinuities only in regions that insufficiently represent original chromatic contrast. All operations are restricted so that they preserve the overall image appearance, lightness range and differences, colour ordering, and spatial details, resulting in perceptually accurate achromatic reproductions of the colour original, as is shown in Figure 1.6(c).

3D Rendering

In 3D rendering, the virtual camera settings are fixed, meaning that not all relevant scene information may be visible. For instance, fine surface details may require more light, which could flatten the shape cues and spatial organization. The challenge is to render all important contrasts in keeping with the scene’s possible animation and camera movement.

The *3D Unsharp Masking* approach enhances local scene contrast by unsharp masking over arbitrary surfaces under any form of illumination. The adaptation of a well-known 2D technique to 3D interactive scenarios is designed to aid viewers in tasks like understanding complex or detailed geometric models, medical visualization and navigation in virtual environments. Its holistic approach enhances the depiction of various visual cues, including gradients from surface shading, surface reflectance, shadows, and highlights, to ease estimation of viewpoint, lighting conditions, shapes of objects and their world-space organization.

Motivated by recent perceptual findings on 3D aspects of the Cornsweet illusion, scene coherent enhancements are created by treating visual cues in terms of their 3D context; doing so has a stronger effect than approaches that operate in a 2D image context and also achieves temporal coherence. The approach is validated with psy-

chophysical experiments showing that the enhanced images are perceived to have better contrast and are preferred over unenhanced originals. The operator runs at real-time rates on a GPU and the effect is easily controlled interactively within the rendering pipeline.

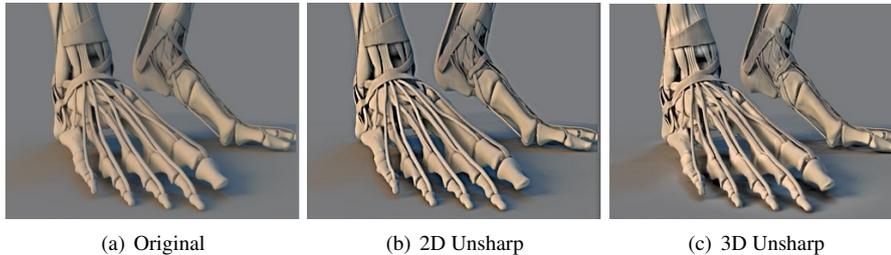


Figure 1.7: A rendering of a 3D model may not contain sufficient contrast to emphasize the geometry (a). For enhancement purposes, the screen space contrasts may be emphasized (b). *3D Unsharp Masking* emphasizes the scene space contrasts (c).

1.3 Artistic Motivation

If we could but paint with the hand what we see with the eye.

– Honoré De Balzac (1799-1850)



Figure 1.8: In *Le Noeud Noir*, Georges Seurat introduces contour effects to enhance overall contrast and emphasize the woman figure against the background.

Can a computer output an image whose appearance closely matches the actual view, or gives a visual impression of being immersed in the scene itself? Images stemming from the most physically correct or mathematically optimal algorithms may not be the most effective for a human viewer. For example, an algorithm that picks each pixel colour so that the widest colour range is used, may not account for the image's overall appearance or may change it in an unnatural way. Likewise, a physically correct rendering of a virtual scene may not depict all of the details due to poor lighting or viewpoint.

Depiction of visual information is the endmost problem of computer graphics. This is a well-known challenge to artists and photographers: how is it possible to represent what the eye sees, from the sun's glow to absence of light, with paints that are slightly brighter than black and much duller than the sun? How can achromatic charcoal and ink depict our colourful surroundings? Over centuries of attempts, painters and photographers have found success not by creating exact reproductions of the light in a scene, but by depicting what they see and what they hope the viewer to see. They intrinsically understand human sight and modify images accordingly by exaggerating contrasts, shifting colours, enhancing detail, lightening shadows and darkening brights. Artists and

photographers must exploit known characteristics of the human visual system (HVS) to produce better imagery, and to push information into the image that otherwise wouldn't be communicated.

The same is true of computer depiction; graphics algorithms can therefore improve by borrowing from artistic techniques and their perceptual motivation [Durand *et al.*, 2002]. This is why, throughout this thesis, the reader will find examples of art that embody motivating techniques or achieve the goal of maintaining contrast appearance despite physical limitations, especially those where specially placed contours increase local contrast. One of many inspirations for adding so-called Cornsweet contours is the drawing shown in Figure 1.8, suggested by Margaret Livingstone [2002]. Despite the limitations of charcoal and paper, Seurat creates the impression of a wider range of tones between the bright and the dark and lifts the figure from the background.

1.4 Novel Contributions

The novel methods presented in the following chapters arise from three publications: *Beyond Tone Mapping: Enhanced Depiction of Tone Mapped HDR Images* [Smith *et al.*, 2006], which enhances the depiction of tone mapped HDR images in a luminance preserving restoration of HDR contrast using colour Cornsweet contours; *Apparent Greyscale: A Simple and Fast Conversion to Perceptually Accurate Images and Video* [Smith *et al.*, 2008], a multiscale technique to convert chromatic lightness to greyscale and restore colour contrast with Cornsweet contours; and *3D Unsharp Masking for Scene Coherent Enhancement* [Ritschel *et al.*, 2008], which performs the scene coherent enhancement of details, shapes, tonal range and spatial organization of a 3D scene within the rendering pipeline.

At the core of each approach is an image processing algorithm that adds Cornsweet contours to alter perceived contrast. The differences and similarities of each algorithm, and the formal explanation of their behaviour, can best be explained in an unsharp masking paradigm. Let us return to the basic definition of unsharp masking, in which an original signal is enhanced by the addition of a *contrast signal* $C(S)$ that is scaled by a *gain factor* usually termed λ [Badamchizadeh and Aghagolzadeh, 2004]. The enhanced signal is defined as

$$\mathcal{U}(S) = S + \lambda C(S) \quad (1.1)$$

where the signal S is usually an image's luminance, but may alternatively be its chromatic component.

The contrast signal, $C(S)$, can be derived in a variety of ways, with the most common technique subtracts a Gaussian blurring of the original from itself, leaving the high-frequencies that were smoothed by the Gaussian, so $C(S) = S - S_\sigma$. The σ parameter controls the bandwidth of the high-frequencies in $C(S)$, which can thus contain only very fine details, or coarser changes over a slightly wider spatial area. In the latter case, local contrast is increased, instead of simply increasing contrast in detailed regions. Regardless of scale, the contrast signal $C(S)$ becomes more clearly depicted in $\mathcal{U}(S)$ than it was in the original S . If $C(S)$ is not present at all in S , then unsharp masking *introduces* a new signal that contains visual information not previously depicted.

The main challenges in designing unsharp masking type methods are: How to define the contrast signal?; In which space to measure the contrast signal (2D image space or 3D scene space)?; Which scale of contrasts the signal should contain (fine detail, larger regions, multiple scales)?; Which channel(s) should be modified (chromatic

or luminance)?; How to control the extent of the enhancement (λ)? The following three algorithms meet each of these challenges by defining the contrast signal, the affected image channel and the manner in which the enhancement is controlled. The high-level diagram shown in Figure 1.9 depicts the flow of information for each. Either the luminance channel, chromatic component or both are used to calculate the contrast signal and the gain value. The contrast is then added either back to the luminance channel, or back to the chromatic component.

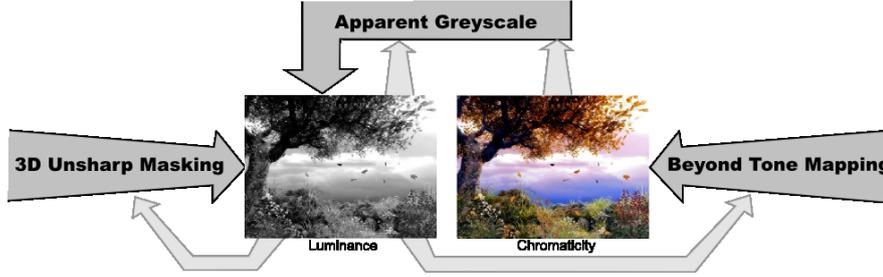


Figure 1.9: Overview of information flow for all three novel approaches. The incoming information is used to calculate the contrast signal and control parameters, which is then output to either the luminance or chromatic channel.

In the *Beyond Tone Mapping* unsharp method¹, λ is simply controlled by the chromatic component values, with $\lambda_{u^*} = u^*$ and $\lambda_{v^*} = v^*$. The contrast signal measures the global and detail contrast that is lost during tone mapping by comparing the original luminance Y to the resulting low-dynamic range luminance y . The polarity of chromatic change in the HDR image's chromatic channel C^* is then incorporated into the contrast signal. Alterations are restricted to the chromatic channels to preserve the tone mapped luminance.

$$\mathcal{U}(I)_{LUV} = [L^*, u^* + \lambda_{u^*} C(Y, y), v^* + \lambda_{v^*} C(Y, y)] \quad (1.2)$$

In the *Apparent Greyscale* method, lost chromatic contrast from the colour image defines λ which is then further controlled by the user. The contrast signal is a set of multiscale high-pass bands from the greyscale image, which gives better control over the width and sharpness of added contrasts. Given a colour original I and its greyscale counterpart G_{L^*} , only the lightness channel is available for manipulation.

$$\mathcal{U}(G)_{LAB} = [G_{L^*} + \lambda(a^*, b^*) C(L^*), a^*, b^*] \quad (1.3)$$

Finally, in *3D Unsharp Masking*, λ is controlled by the user and then automatically adapted to the amount of texture details at each location. The contrast signal is calculated within the synthetic 3D world, and then alters the lightness channel. The chromatic channels are appropriately adjusted to prevent hue shifting. This scene space approach overcomes the temporal problems of screen space approaches.

$$\mathcal{U}(S)_{LAB} = [L^* + \lambda C(S_{L^*}), a^*, b^*] \quad (1.4)$$

The main contributions of these methods are as follows:

¹The distortion method does not fit into the unsharp masking framework.

Luminance Preserving Contrast Adjustment An image processing technique to add chromatic contrast without changing the explicitly determined lightness of a tone mapped image.

Perceptually Accurate Greyscale A perceptually validated conversion of colour images and video to their apparent lightness using models from colour theory.

Video to Greyscale The first greyscale conversion algorithm suitable for video that properly treats isoluminant colours and preserves the appearance of the colour original.

Scene Coherent Enhancement The first enhancement technique to address all lightness contrasts and the first suited for complex geometry, artist-directed lighting, multiple objects and animated scenes.

3D Cornsweet Illusion A physically correct method for generating 3D Cornsweet illusions.

1.5 Thesis Organization

This thesis begins with an overview of contrast perception, contrast enhancement and contour illusions. Chapter 2 provides a background on vision and contrast perception, then moves to contrast measures for complex images. Section 2.3 gives an extended overview of colour contrast and chromatic lightness. After reviewing contrast perception and its related metrics, people's preference for greater contrast and more colour saturation in images is discussed. To conclude, contrast enhancement algorithms are outlined, including an in-depth review of unsharp masking techniques. Chapter 3 presents contour brightness illusions and explains their measured impact on perceived contrast. Emphasis is placed on the Cornsweet illusion and its cause, the Cornsweet contour. It also presents findings on the effect of chromatic brightness illusions and illusions in 3D. The chapter concludes in a discussion on several perceptual theories that account for the contour effects, and a statement on the relationship between unsharp masking and the Cornsweet illusion.

In the middle three chapters, focus shifts to three different topics in Computer Graphics, each related to contrast enhancement. Chapter 4 presents tone mapping operators and techniques for preserving the impression of wider dynamic range. Section 4.3 presents the novel *Beyond Tone Mapping* method for restoring contrast using colour contours. Chapter 5 motivates the problem of greyscale conversion and provides a wide survey of recent approaches. Section 5.3 presents the novel *Apparent Greyscale* perceptually accurate conversion method. Applications of this approach and comparisons to other conversion methods are given in Section 5.4. Chapter 6 presents the problem of enhancing the depiction of 3D scenes and various recent techniques. Section 6.3 details the novel *3d Unsharp Masking* that adapts the 2D algorithm to a 3D context. Section 6.4 gives an assortment of applications and shows how well the novel use of unsharp masking within a scene creates more effective and informative renderings. An artistic motivation is presented in each of these three chapters, as well as a discussion on the use of Cornsweet contours to solve each problem compared to other existing approaches.

Chapter 7 concludes the thesis by recapping all that has been presented and highlighting the major contributions and future directions of this work.

Chapter 2

Contrast Perception and Enhancement

This chapter presents a background on contrast perception, and discusses physiological and computational aspects of contrast. The information collected here identifies contrast as a fundamental aspect of visual perception and builds a basic groundwork for measuring contrast in images, which is needed to understand the methods presented in Chapters 4, 5 and 6.

To begin, Section 2.1 is a short overview on how imagery is processed by the human visual system, the role played by contrast and the effects of apparent contrast. In Section 2.2, various measures of contrast are defined, including a multi-scale contrast-representation of images.

A detailed description of colour contrast and colour brightness perception is provided in Section 2.3. This is included because it is necessary for understanding how contrast is measured in colour images. Additionally, Section 2.3.3 provides an introduction to chromatic lightness, which is an integral aspect of the greyscale conversion method presented in Chapter 5.

In photography and publishing, higher contrast images have a stronger visual impact and appreciation. Due to this preference for higher contrast, many image processing techniques increase contrast to enhance or otherwise improve the original image. Section 2.4 provides a context for the methods in Chapters 4, 5 and 6, by explaining contrast enhancement techniques, particularly *unsharp masking* approaches. Section 2.3 details findings that people not only perceive enhanced images as having a higher quality, but have been shown to prefer photos whose contrast has been increased.

2.1 Seeing Contrast

Conceptually, contrast is the spatial variation of a visual stimulus. In cognitive respects, contrast as a perceptual quality arising from differences in brightness is a key to visual understanding: how we distinguish one visual element from another, the shape of an object, its texture and details. The term contrast is intentionally vague, since it has both physical and perceptual definitions, and depends on whether the stimulus is a real scene, a photograph or a simplified test image. Contrast can be defined objectively as a difference between the luminance reflected by two regions or surfaces. Perceptually, it is the difference in perceived brightness, which makes it dependent on attributes such as surround, chroma, local spatial structures and environmental aspects.

The following four sections show how and why vision is heavily based on local contrast, and how human sensitivity to contrast is a factor of average illumination and spatial frequency. The theories and science are presented to give a general context to terminology, and so the reader may understand the motivation behind the novel contributions detailed in following chapters. For a complete picture of the physiological and neurological process of seeing, the reader is referred to the classic text by Stephen E. Palmer entitled *Vision Science* [1999], or the entertaining anecdotal book entitled *Vision and Art: The Biology of Seeing* by Margaret Livingstone [2002].

2.1.1 Lightness constancy

There are several terms to describe different aspects of seeing light. To begin, light is reflected off a surface to an eye. The amount of physical light is termed *luminance*.

Luminance The physical amount of light reflected off a surface or emitted from a light source in a particular direction, measured in cd/m^2 and denoted as Y .

Through various levels of physiological and neurological processing, luminance is perceived by a viewer as *lightness*:

Lightness The perceptual response of a human viewer to luminance and is defined by Hunt as “the brightness of objects relative to that of a similarly illuminated white”. Denoted L^* (“L-star”), it is defined in terms of luminance Y and reference white luminance Y_n .

$$L^* = 116(Y/Y_n)^{1/3} - 16$$

It can also be observed that the definition is relative to a fixed white level. This relativity accounts for the phenomenon of *lightness constancy*. Lightness constancy is convincingly illustrated by Edward Adelson’s now famous example, reproduced in Figure 2.1. The squares marked A & B have identical grey values, however B is perceived as much brighter. It is as though the effect of the cylinder’s shadow, which darkens the white square (B), is discounted so that it appears consistent with the non-shadowed white squares.

Lightness Constancy The lightness (perceived luminance) of a surface is consistent under differences in illumination and viewpoint. In other words, the impact of shadows/lighting are disregarded.

Lightness is not the ultimate percept, however. There are other factors that affect the way in which lightness is perceived. The term *brightness* is a non-relative way to describe the perception of light.

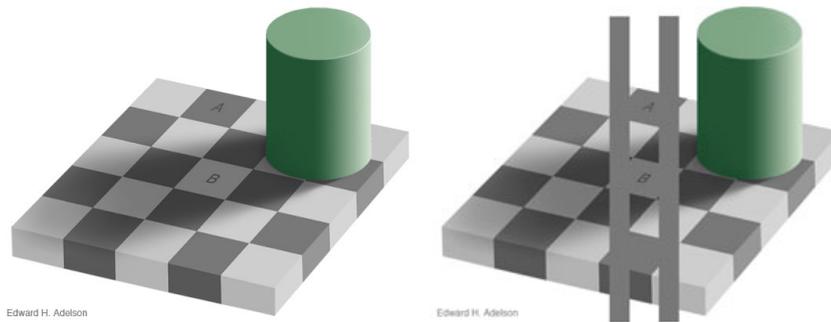


Figure 2.1: Lightness constancy illustrated by Edward Adelson's now famous example. A & B have identical grey values, however, the one in shadow is perceived as much brighter because the illumination of the scene is discounted.

Brightness Attributes of a visual sensation according to which an area appears to emit more or less light [Fairchild, 2005].

The term brightness encompasses the effect of lightness constancy, but includes many more factors. It is a term to describe the perception of light, including the possibly unexplained effects of spatial organization, context, and environment. In particular, brightness is affected by local luminance ratios at edges, as detailed in Section 2.1.2

2.1.2 Spatial Vision

To explain lightness constancy, German physiologist Ewald Hering observed that under different illumination, absolute luminance levels change, but the relative luminance of one region compared to a neighbouring region remains nearly unchanged. Thus, he proposed that brightness is dependant on the contrast between regions. Hans Wallach proposed a slightly more specific theory by stating that brightness is determined by luminance ratios, thus specifying the way in which contrast is measured.

So brightness is strongly influenced by contrast, or is at least relative to surrounding areas. But how is this contrast evaluated? What spatial processing is involved that assesses the luminance ratios? As evidenced by the Cornsweet illusion depicted in Figure 1.1, local luminance ratios that occur at edges are more influential than luminance ratios of distant regions. The Retinex theory, formulated by Land and McCann, models lightness perception as a global integration over local contrast at edges, or contrast caused by spatially adjacent regions. This theory does indeed predict the perceptual effect of lightness constancy for smoothly changing illumination without discontinuities such as shadows, and also correctly predicts the effect of the Cornsweet illusion.

Because local contrast plays a very important role in visual processing, there has been much research on the spatial aspects of vision. Physiological experimentation found evidence of spatial processing between neighbouring regions of the retina. Modelling the visual nervous system as a neural network, the popular laterally inhibitory organization between layers of cells can describe the way in which local contrasts are perceived.

Lateral inhibition means that the incoming stimulus to a photoreceptor then excites a ganglion cell in the next level of the retina, while at the same time inhibiting the responses of nearby cells. Physiological experiments have shown that retinal gan-

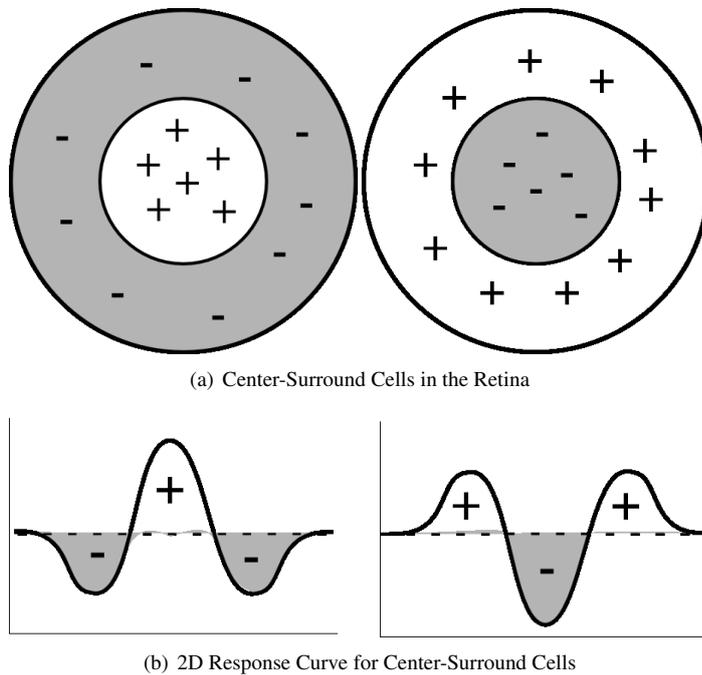


Figure 2.2: Lateral inhibition describing the receptive fields of center-surround cell in the retinal ganglion. Modelled after Figure 4.3.5 in [Palmer, 1999]

glion cells come in two types, depending on the receptive field that combines incoming impulses from a region of photoreceptors. These are the on-center/off-surround and the off-center/on-surround cells, generally called center-surround cells. Diagrammatic versions of the cells' receptive fields are shown in Figure 2.2, along with the response profiles for a 2D signal. The response of retinal ganglion cells depends on the spatial interaction of excited photoreceptors; their response is derivative in nature as they respond to changes in contrast.

The transition between two contrasting adjacent regions creates a visible *edge*, which may be gradual and is distinguished here as a *contour*. These are areas of local contrast, to which the visual system produces strong responses. The ring-like receptive fields of the center-surround cells function to clarify the incoming luminance information by emphasizing local contrasts which results in visible contours.

The behaviour of spatial vision has been investigated with computational models that emulate the behaviour of the visual system. Since spatial transitions can be characterized in several ways, the visual system's response has been modelled by several different types of operators. The second-order derivative operator is closest to the Retinex theory and the center-surround principles. As can be observed when comparing Figures 2.2(b) and 2.9, the receptive field of the retinal ganglion cells matches the second-order derivative Laplacian operator.

A second-order derivative operator, such as the Laplacian or the difference of Gaussians, computationally model the localized behaviour of the visual system, as in Figures 2.3(a) and 2.3(b). Its response to an image measures the steepness of local contrasts and its zero-crossings identify where edges occur as shown in Figure 2.3(c). Notice that the second-order derivative image is roughly uniform where there is no contrast.

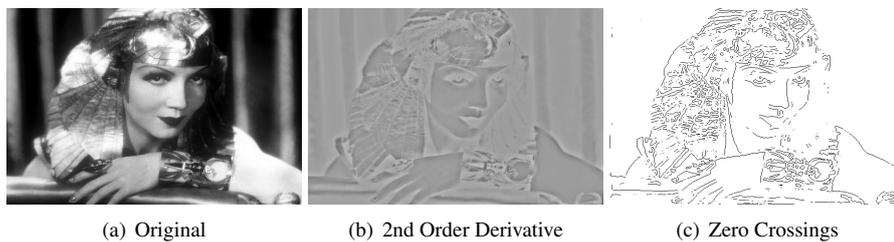


Figure 2.3: A second-order operator simulates the derivative nature of vision and contrast processing. The contours of the second-derivative identify the changes in contrast, and edges occur at their zero-crossings.

2.1.3 Spatial Frequency and Contrast Sensitivity

How big must a luminance difference be in order to be visible? The *Weber law* tells us that visual contrast sensitivity is constant in proportion to mean local luminance in the bright range, then decreases as illumination descends to dark. Now, how is contrast perception affected by the fineness or coarseness of the frequency of luminance variation?

One theory of visual image processing is that images are interpreted by psychophysical channels at various spatial frequencies. These theoretical channels create a theory of vision which is supported by evidence that low-level cortex cells in visual systems acts as spatial filters by performing a *local spatial frequency analysis* [Palmer, 1999].

These psychophysical channels, that for instance contain the low-frequency information and high-frequency information of a complex image, may be processed separately. As shown with the the band-limited images in Figure 2.4, the low-frequencies form a base image that communicates the more global image features, like the large-scale transitions from dark to light, Figure 2.4(a); the high-frequencies communicate finer details, like texture and local contrasts, Figure 2.4(c). This theory has been modelled by a scale-space representation that splits an image into several band-pass channels believing this is the way HVS splits imagery for understanding [Valois and Valois, 1990].

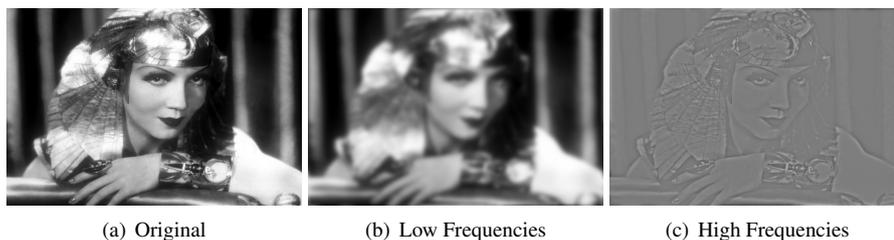


Figure 2.4: The spatial frequency components of an image, which according to vision theory, are processed separately by spatial psychophysical channels. These are sometimes known as the base (b) and detail (c) layers of an image.

Much effort has gone into measuring the effect of different spatial frequencies on perceived contrast. Using sinusoid grating functions, the *Contrast Sensitivity Func-*

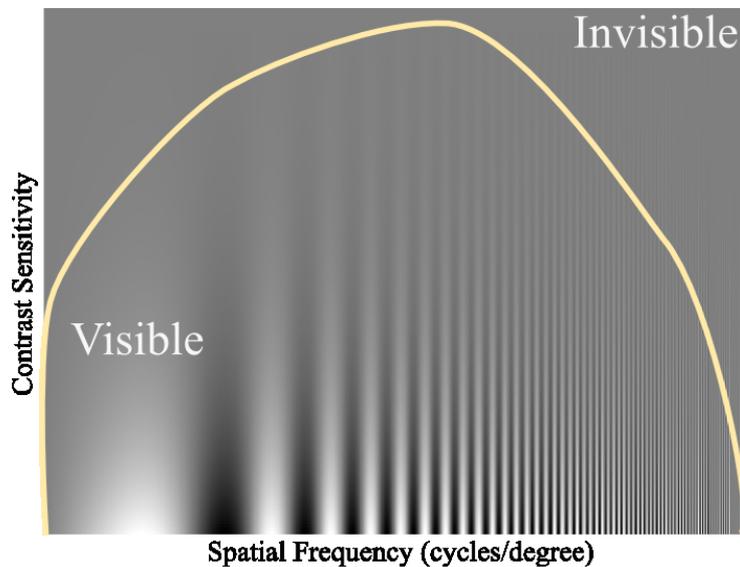


Figure 2.5: A diagram representing the shape and values of humans' Contrast Sensitivity Function (CSF). Contrast sensitivity is reduced at high-frequencies.

tion (CSF) defines the perceptual response to simple stimuli at a broad range of spatial frequencies and increasing contrast magnitudes, shown here in Figure 2.5. Contrast sensitivity measures the lowest contrast at which there is a perceived difference between a grating image and a uniform grey background. The CSF shows that humans have better sensitivity at low-frequencies, when the gratings are wider. There are physiological reasons for the CSF, for instance, limited density of photoreceptors in the retina can explain the fall-off of sensitivity for high-frequencies [Palmer, 1999].

2.1.4 Apparent Contrast

Why is apparent contrast more than a simple difference between incoming luminances? To begin, the surrounding luminance (or the illumination) has a marked effect on perceived brightness, and thus on the apparent contrast between a stimulus patch and its surround. This phenomenon is named *simultaneous contrast*. Although of the same luminance, the brightness of each center square in Figure 2.6 lightens as its surround darkens, and vice versa. Thus, brightness is relative, depending on a background luminance. This phenomenon also suggests that the human eye is more sensitive to contrast than absolute luminance.

The effect of surround on apparent contrast has even been used to create visual illusions, as shown in Figure 2.7(a). While the illusion is a fun surprise, and serves to convince us not to believe exactly what we see, it can also be put to use as an artistic technique. The effect of simultaneous contrast was understood as early as the seventeenth century by Dutch still life painter Willem Claeszoon Heda, who created a diagonal intensity gradient across the background, to make the bright foreground regions seem brighter and the dark regions darker, Figure 2.7(b). In doing so, he creates an impression of greater dynamic range, more like that of the real world than of a painting. The approach makes the left hand side objects appear darker compared to



Figure 2.6: An example of the simultaneous contrast effect: although of the same luminance, the brightness of each center square increases as its surround darkens.

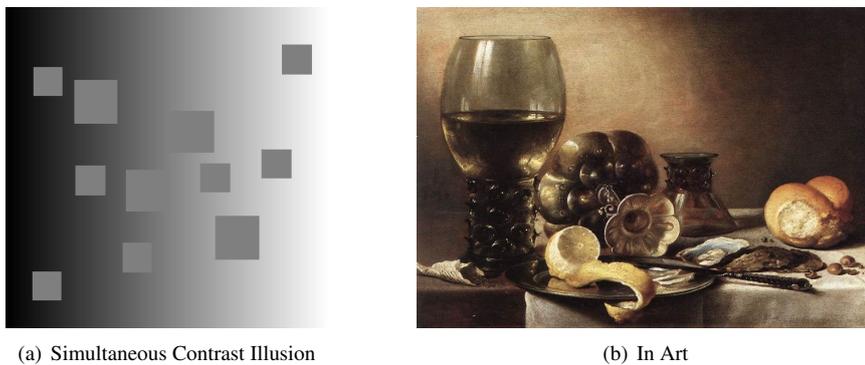


Figure 2.7: An illusion where a gradient background makes equiluminant squares appear to have different brightness levels (a). Dutch painter Willem Claeszoon Heda uses a simultaneous contrast effect to make the wine glass appear darker, and the lemon and bread roll brighter.

their brighter surround, and the right hand side bread and cloth appear much brighter due to its dark surround.

Local or structural effects also expose the complex relationship between luminance and perceived brightness. For instance, the Cornsweet Illusion presented in Figure 1.1(a) is a very convincing example. Another related phenomenon in Ernst Mach's famous *Mach Bands*, which name the illusory light and dark peaks being perceived on either edge of a gradually sloping contour. This phenomenon, named *Mach Banding*, is shown in Figure 3.3(b) and is another excellent demonstration of how our eyes respond to luminance contrast in sometimes surprising ways [Ratliff, 1965]. Chapter 3 is dedicated to local contrast illusions such as these so-called contour effects, and discusses their perceptual explanations.

2.2 Measuring Contrast

Contrast is difficult to measure and to define, especially in natural images. By convention, most metrics are defined either for grating image stimuli, such as those used to study human vision and pattern perception (Section 2.1.2) or on simple uniform patches overlaying uniform backgrounds. Grating images are made up of alternating dark and light bars; their contrast is the intensity difference between those bars. A

greater difference is said to have higher contrast, and as the difference diminishes the bars become indistinguishable, leading to subthreshold contrast.

The most simple measure of contrast is the ratio of maximum to minimum luminance. Note that instead of measuring contrast of luminance L , one could also measure the contrast of perceived lightness L^* .

$$C_{Basic} = \frac{L_{max}}{L_{min}}$$

To account for a wider range of luminances, like those present in the real world, and humans' greater sensitivity to changes in the higher range of luminance, the logarithmic ratio has been used to define contrast. Especially for HDR images, the logarithmic ratio gives a better measure of visible contrast, and is defined as:

$$C_{LogRatio} = \log 10 \frac{L_{max}}{L_{min}} \quad (2.1)$$

For simple grating images with maximum and minimum luminances L_{max} and L_{min} , contrast is measured with the *Michelson formula*

$$C_{Michelson} = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}$$

A better contrast measure for a simple target and its surround is the *Weber fraction*. It is measured for background L and a target patch that is ΔL brighter or darker than the background. However, this normalized difference measure works best when ΔL is not large, and when there is a single target on the background. The Weber fraction is generally used for measuring the contrast of patches of different luminance, and is defined as:

$$C_{Weber} = \frac{\Delta L}{L} \quad (2.2)$$

2.2.1 Local Contrast

In natural images, contrast varies spatially over the depicted scene. There may be both large flat areas free of contrast (i.e. a clear sky) and areas with high contrast (i.e. a field of sunflowers). The contrast measures given in the previous section are suitable for simple stimuli, but may not capture the appearance of contrast in a complex image. A single measure applied over the whole image is susceptible to highlights and shadows and outlier luminance values. Due to two greatly differing regions, a high contrast measure may occur even when overall perceived contrast seems very low.

A simple measure could be made local by applying it to each pixel and its surrounding neighbourhood. A more perception related metric incorporates findings in vision theory (Section 2.1) that show local contrast is related to an image's local gradient [Valois and Valois, 1990]. Local contrast is modeled as an abrupt change in intensity (i.e. luminance), and the maximum of that change is said to be an *edge*. Thus edges represent locations of local contrast, and are found by *edge detection*.

So, the magnitude and orientation of edges in an image are a measure of its local contrast. Sharp contrasts (those that produce obvious edges) are measured by calculating the rate of change of intensity, otherwise known as the *intensity gradient*. These first-derivative methods measure the amount of difference between pixels, and the direction of greatest change.

For a host of reasons, including image understanding, computer vision and robotics, edge detection is a fundamental task in image processing, and there are various different algorithms to perform the task. For instance, given the intensity profile of an edge in Figure 2.8(a), the peak in the first-order gradient, shown in Figure 2.8(b), detects edges in the image and thus identifies intensity difference. Approximate gradient calculation can be done efficiently with convolution kernels, like the Roberts, Prewitt or Sobel kernels [Valois and Valois, 1990].

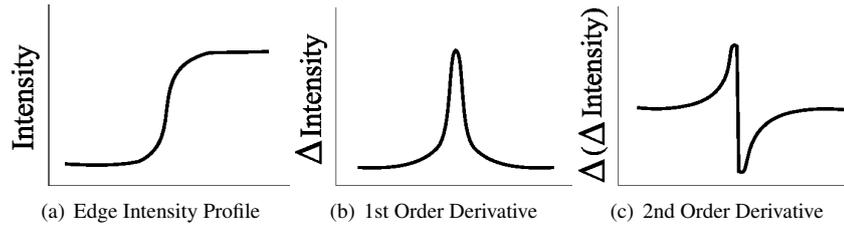


Figure 2.8: An edge signal and its derivatives modeled after the Figure 4.3.4. in [Palmer, 1999].

However, contrast is more often a gradual transition between light and dark, not a perfectly abrupt change. Thus, local contrast is less characterized as an edge, and more by a *contour* defined by where the gradient magnitude rises, peaks and then falls. Where the ideal edge separating two intensities occurs at the peak, the whole transition is characterized as a contour that contains information about the spatial properties of the transition, including the abruptness of the contrast.

To better capture the contrast of gradual intensity transitions, second-order derivative filters measure contrast as the rate of change in the intensity gradient (the divergence of the gradient). For instance, the second-order derivative of an edge given in Figure 2.8(c) characterizes its contrast, and its optimal location is found at the 2nd derivative's maxima and minima, the peaks of the intensity gradient. The second gradient is positive on the dark side of the edge, and negative on the bright side. Peaks occur where the second derivative is zero, but since images are discrete, they occur at the zero crossings - where one value is negative and the neighbouring value is positive.

Since the gradient of the image (its first derivative) defines a vector field, the image's second-order derivative is the divergence of the gradient. Thus, the second derivative can be measured by the *Laplacian* of the image. This for instance is done in the Marr-Hildreth edge detection method [Valois and Valois, 1990]. However, the Laplacian operator has been shown to be very sensitive to noise. For this reason, the *Laplacian of a Gaussian* (LoG) operator is used, denoted as ∇^2 . Before applying the Laplacian, this operator removes noise with a Gaussian smoothing operator with standard deviation σ (gives the blur radius), defined as:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.3)$$

To increase efficiency, these convolutions are replaced by a single convolution kernel created by differentiating the Gaussian blur. The blur radius now controls which contrasts will be discarded before measurement. Differentiating the Gaussian results

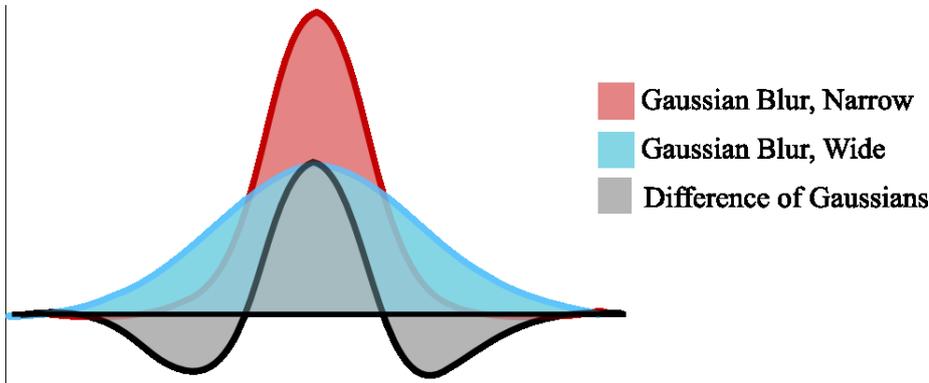


Figure 2.9: The difference of two Gaussians approximates the Laplacian of a Gaussian (LoG operator). In this example, the wider Gaussian is subtracted from the narrower.

in the definition of the LoG, $\nabla^2 G(x, y)$, as:

$$LoG(x, y) = -\frac{1}{\pi\sigma^4} \left(1 - \frac{(x^2 + y^2)}{2\pi\sigma^2}\right) e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (2.4)$$

Another second-derivative edge detection operator is based on the fact that the LoG can be approximated as the difference of two Gaussians. This is true because the LoG is the derivative of a Gaussian with respect to $2\sigma^2$ of a Gaussian, and is called the *Difference of Gaussians* (the DoG operator). DoG is defined as one Gaussian blurred image minus another blurred with a slightly smaller radius. The DoG operator also acts as a band-pass filter resulting in a quasi-band limited image that measures contrasts of a specific scale. Thus, the measured contrast are more blurred or more sharp, depending on the size of Gaussians used in the calculation.

2.2.2 Local Band-limited Contrast

The local contrast measures from Section 2.2.1 are better at capturing the spatially-varying natural of contrast, however they do not address the band-limiting nature of vision that also affects contrast in an image. For instance, very small-scale local contrasts are named details and relate to the impression of texture in a scene, whereas slightly larger-scale contrast are those that influence geometry contrast and differences between shapes. To process images in a way that interprets differences in scale, pyramid methods have proven useful, if not integral [Adelson and Ogden, 1984]. Derivative-based local contrast measures presented in Section 2.2.1 were shown to better approximate the derivative response of the eye's receptive fields, as described in Section 2.1.2. The Laplacian of Gaussian (LoG) and Difference of Gaussian (DoG) approaches both include methods for adjusting the scale of the resulting contrast measures. To measure local contrast on various different levels, these derivative approaches can be extended to multi-scale measures using pyramids or scale space. For accuracy, measurements should be done in scale-space [Lindeberg, 1994]. However, pyramids are still worthy, as they are computationally efficient approximations of scale-space.

The first and simplest measure is the Laplacian pyramid [Burt and Adelson, 1983]. An image is decomposed into different frequency bands with the LoG (Equation eq:log)

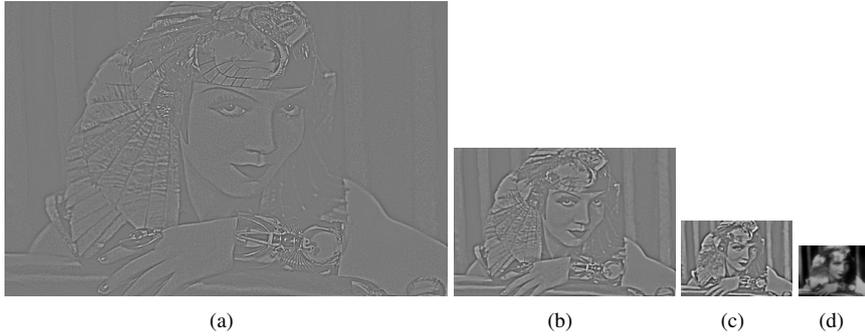


Figure 2.10: An efficient approximation of space-space image representation: the Laplacian pyramid by [Burt and Adelson, 1983].

which is calculated at each band using DoG. Only the contrasts contained in the LoG images are stored, plus the final image which contains the remaining low-frequencies. Each level of the pyramid is a successively smaller image that is not exactly the Laplacian, but rather, the negative of the Laplacian. This is because a wider Gaussian is subtracted from a smaller Gaussian, or in the first level, from the original un-blurred image. A visualization of the DoG operator is shown in Figure 2.13 Each level carries the local contrasts at a different range of frequencies and a base level which collects the lowest remaining frequencies. This image-representation both mimics the local spatial processing and the band-limiting nature of human vision, described in Section 2.1. More formally, the Laplacian pyramid that decomposes an image into n bandpass images h_i and a single lowpass image l . For image I , where each h_i measures its local contrast at scale i (lowest at n), the decomposition is:

$$I = \sum_{i=0}^{n-1} h_i(I) + l \quad (2.5)$$

A recent pyramid structure by Peli captures frequency-dependent properties of spatial vision with his local band-limited contrast measure for dealing with complex images [Peli, 2000]. In his work, contrast is measured by filtering and comparing increments or decrements from the local background intensity. Contrast is defined for each frequency band as the ratio of intensity at a pixel and the average luminance of its neighbourhood. Given a pyramidal image, the structure is defined as:

$$I = l_0 + \sum_{i=1}^{n-1} a_i, \quad l_i = l_0 + \sum_{j=1}^{i-1} a_j \quad (2.6)$$

where a_i is the bandpass-filtered image, l_0 is the lowest band, and l_i is a corresponding local mean luminance image containing all frequencies below the band. Thus, at the band of spatial frequencies in a_i , contrast is defined locally as:

$$C_{Peli}(x,y) = \frac{a_i(x,y)}{l_i(x,y)} \quad (2.7)$$

If there is a need for more accuracy in the contrast measures, then they should be related to the CSF from Section 2.1.3. One approach is to multiply each band of the

multi-scale image representation by a factor taken from the CSF according to the band's mid-frequency. Or, to model the frequency dependence of the CSF, contrast can be measured in the frequency domain instead of the spatial domain by converting the image to Fourier space [Gonzalez and Woods, 2002]. Another perception-based contrast metric is the "Global Contrast Factor" that computes contrast at different resolutions to define a global measure that corresponds better to a human's perception [Matkovic *et al.*, 2005]. Mantiuk *et al.* describe a contrast measure for HDR images that uses the perceptual "Just Noticeable Difference" by transforming contrast magnitudes through a transducer function to model their threshold or super-threshold behaviour [Mantiuk *et al.*, 2006b]. For HDR stimulus, the use of logarithmic ratio of luminance accounts for the logarithmic behaviour of brightness perception. It also gives a better measure for LDR and HDR images, and makes comparison between their contrasts slightly better. This logarithmic ratio is returned to in Section 4.3.3.

2.3 Chromatic Contrast

"In visual perception a color is almost never seen as it really is – as it physically is. This face makes color the most relative medium in art."

– Josef Albers (1888-1967) [1975]

So far, contrast has been treated as an achromatic percept. However, light is composed of different wavelengths that produce colour sensations in a human viewer. The physiological workings and theory of colour vision itself is very interesting, but beyond the scope of this thesis. The interested reader is referred again to [Palmer, 1999] for a detailed account of color vision, and for a less formal treatment to [Livingstone, 2002]. One basic theory is that luminance and colour are processed separately in the visual system, and that luminance communicates the *where* (shapes, lighting, depth) and colour communicates the *what* (texture, material). These two sets of information are then combined at the cognitive level to produce the final percept.

Measuring the contrast of coloured stimuli can be done simply by treating each colour channel (i.e. R, G and B) as separate achromatic signals. However, there are more accurate ways to measure colour contrast, as described in Section 2.3.2. Finally, as there are colour effects that change the perceived lightness of a colour, it may be useful to measure the contrast of the predicted chromatic lightness, a concept introduced in Section 2.3.4.

2.3.1 Colour Model Background

In order to discuss colour, we must move into the realm where colour exists - that is, in the observer's mind. Despite the elusive nature of colour, colour theorists and scientists know that a physical stimulus produces a perceptual response that we name "colour", or alternately, an achromatic response we name *perceived lightness* or *brightness*. Colour appearance models take on the complex task of predicting a human viewer's perceptual response to colour stimulus and of defining measures of colour. The models are fit to data from careful psychophysical experiments that use simple stimuli of isolated colours. Complex colour appearance models and their associated metrics are based on additional data with the goal of predicting a viewer's perceptual response to the sight of a colour, even under drastically changed lighting conditions [Fairchild, 2005].

Theoretically, the models provide a system for describing colour as a combination of numbers. For instance, in the common *RGB* colour model, colours are defined according to their three colours primaries: red *R*, green *G* and blue *B*. When the specific use of the model is defined by giving the viewing conditions or by providing a mapping to a reference absolute colour space, a new colour space is defined.

In the 1924 Commission Internationale de l'Éclairage (CIE) *RGB* space, colours were defined according to a standard observer's colour matching task of choosing the amounts of pure red, green and blue that when mixed, would match a stimulus colour. The *RGB* model is actually used in most digital image representations by specifying the colours in one of the many existing *RGB* colour spaces, like *sRGB* or Adobe *RGB*. However, *RGB* spaces are not always suitable for image processing, since the differences in the space do not correspond to perceived colour differences, and some negative amounts were required to match all visible colours. Additionally, the mixing of primaries is not as intuitive as adjusting colour along perceptual axes, like hue, saturation or value.

2.3.2 CIE $L^*a^*b^*$ and $L^*u^*v^*$

A variety of colour spaces based on human perception have been endorsed by the Commission Internationale de l'Éclairage (CIE). To eliminate negative primary values and to create one primary that corresponds to luminance or lightness, CIE endorsed another set of primaries known at the 1931 CIE *XYZ* tristimulus values. These primaries roughly align to the high, mid and low wavelengths of light to which eyes are sensitive.

The standard CIE colour space from 1935 is a colourimetry system that defines colour tristimulus values based on colour matching experiments. Of the three *XYZ* primaries, *Y* singularly defines the luminance of a stimulus. The colour space is illuminant dependent, meaning that the colour coordinates are relative to a specific white point, denoted as X_n, Y_n, Z_n . For most standard purposes the CIE standard illuminant D65 is used, since it roughly corresponds to average daylight.

Stemming from the tristimulus values, CIE 1976 $L^*a^*b^*$ (*CIELAB*) and $L^*u^*v^*$ (*CIELUV*) are more intuitive 3d spaces. They are easily derived from the *XYZ* space, but have the advantage of being more perceptually uniform: measured differences in the colour space match better to the visible difference between colours. The three axes of both the *CIELAB* and *CIELUV* colour spaces approximate perceived lightness L^* , chroma C^* and hue angle H^* . Recall from Section 2.1.4, that the first component, L^* , quantifies the relative perceptual response of a human viewer to luminance and is defined by Hunt as “the brightness of objects relative to that of a similarly illuminated white”.

Lightness The brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting [Fairchild, 2005].

Mathematically, lightness is defined in terms of luminance *Y* and reference white luminance Y_n (when $Y/Y_n > (6/29)^3$):

$$L^* = 116(Y/Y_n)^{1/3} - 16 \quad (2.8)$$

CIELAB and *CIELUV* are both colour spaces with L^* representing a lightness and the other channels representing decorrelated opponent chromaticities. The chromatic dimensions correspond to the red-green and yellow-blue opponent colour pairs. They

are simple, but widely used colour spaces. While both have very similar properties, their difference lies in their chromatic calculations.

CIELAB defines its two colour components a^* and b^* as a multiplicative normalization of X, Y and Z by the reference white:

$$a^* = 500[f(X/X_n) - f(Y/Y_n)] \quad (2.9)$$

$$b^* = 200[f(Y/Y_n) - f(Z/Z_n)] \quad (2.10)$$

$$f(\omega) = \begin{cases} \omega^{1/3} & \omega > 0.008856 \\ 7.787\omega + 16/116 & \omega \leq 0.008856 \end{cases}$$

Whereas, *CIELUV* is a subtractive shift from the reference white whose chromatic components u^* and v^* are defined as:

$$u^* = 13L^*(u' - u'_n) \quad (2.11)$$

$$v^* = 13L^*(v' - v'_n) \quad (2.12)$$

$$u' = 4X/(X + 15Y + 3Z) \quad v' = 9Y/(X + 15Y + 3Z) \quad (2.13)$$

It is sometimes simpler to consider these chromatic dimensions in terms of their colour-related correlates hue h , which is the hue angle over a colour disk, and chroma C^* , its perceived colourfulness. These correlates define the cylindrical versions of *CIELAB* and *CIELUV*, LCH_{ab} and LCH_{uv} respectively.

$$h_{ab} = \arctan(b^*/a^*) \quad h_{uv} = \arctan(v^*/u^*) \quad (2.14)$$

$$C_{ab}^* = \sqrt{a^{*2} + b^{*2}} \quad C_{uv}^* = \sqrt{u^{*2} + v^{*2}} \quad (2.15)$$

Saturation s is defined in several ways, but in general describes its perceived intensity with respect to its lightness. Thus, although not recognized by the CIE, it can be estimated as chroma normalized by lightness, as follows:

$$s_{ab} = \frac{\sqrt{a^{*2} + b^{*2}}}{L^*} \quad s_{uv} = 13\sqrt{u^{*2} + v^{*2}} \quad (2.16)$$

Colour difference, or chromatic contrast, is defined as the Euclidean distance between two colours in a 3D colour space. Colour difference is illuminant dependent, and different standard measures exist, each of which models colour distance slightly differently. The simplest recommended difference measure is CIE 1976 ΔE_{ab}^* , which for two colours is

$$\Delta E_{ab}^* = \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}} \quad (2.17)$$

Fairchild describes other more exact and complex ΔE^* measures in [2005].

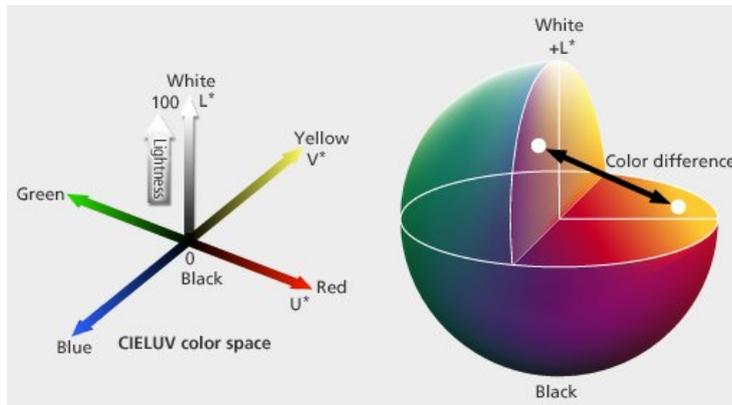


Figure 2.11: A visualization of the CIELUV colour space and a colour difference (from Wikipedia.org).

| Common Term | Technical Term | Notation |
|-----------------------------|------------------------------|--------------------------|
| Brightness | Luminance | Y |
| Chromatic component | CIE 1976 chromatic primaries | X, Z |
| Lightness | CIE 1976 lightness | L^* |
| Hue and saturation | Chromaticity | u', v' |
| Relative hue and saturation | CIE 1976 | u^*, v^* or a^*, b^* |
| Chroma | CIE 1976 chroma | C_{uv}^* or C_{ab}^* |
| Hue | CIE 1976 hue-angle | h_{uv} or h_{ab} |
| Saturation | CIE 1976 saturation | s_{uv} |
| Colour difference | Delta E | ΔE_{ab}^* |

Table 2.1: Important colour terms, modified from The Reproduction of Color, [Hunt, 1995].

2.3.3 The Helmholtz-Kohlrausch Effect

While a colour's luminance is the dominant contributor to lightness perception, colour phenomena is another important factor. Colour phenomena capture the fact that isoluminant colour stimuli have different perceived lightness, or more generally, that the chromatic component contributes to perceived lightness, and that this contribution varies according to both hue and luminance. This phenomenon is characterized by the Helmholtz-Kohlrausch effect, where given two isoluminant colours, the more colourful sample appears brighter.

Helmholtz-Kohlrausch Effect A chromatic stimulus with the same luminance as a white reference stimulus will appear brighter than the reference [Nayatani, 1997].

There are two experimental approaches for measuring the H-K effect: the Variable-Achromatic-Colour (VAC) approach, in which an achromatic sample's luminance is adjusted to match a colour stimulus; and the Variable-Chromatic-Colour (VCC) approach, in which the chromatic content of a colour stimulus is adjusted until its brightness matches a given grey stimulus [Nayatani, 1998]. VAC is more common and was used in the seminal 1954 Sanders-Wyszecki study, and again in Wyszecki's later 1964

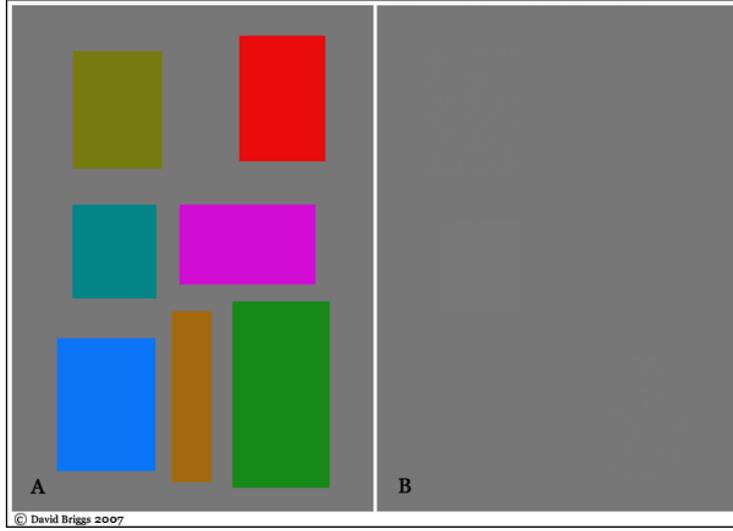


Figure 2.12: This image is isoluminant, however the colour patches have a clear contrast against the achromatic background. This phenomenon is named the Helmholtz-Kohlrausch effect (H-K effect). From David Briggs, “The Dimensions of Colour” <http://huevaluechroma.com/>.

and 1967 studies [Wyszecki, 1967].

2.3.4 Survey of Chromatic Lightness Predictors

The H-K phenomenon is predicted by a *chromatic lightness term* that corrects L^* based on the colour’s chromatic component. Fairchild’s *CIELAB* chromatic lightness metric L^{**} is fit to Wyszecki 1967 data. Chroma C^* measures colourfulness and a sinusoidal curve predicts the H-K effect’s decreased impact at yellow hues and its strongest effect at blues. It is defined as [Fairchild and Pirrotta, 1991]:

$$L^{**} = L^* + (2.5 - 0.025L^*) \left(0.116 \left| \sin \left(\frac{H^* - 90}{2} \right) \right| + 0.085 \right) C^* \quad (2.18)$$

Nayatani defines chromatic lightness metrics $L_{N_{VAC}}^*$ and $L_{N_{VCC}}^*$, for each experimental approach, based in CIELUV¹ [1997]. A quantitative difference between them is that $L_{N_{VCC}}^*$ is twice as strong $L_{N_{VAC}}^*$ in log space. For each method, chromatic object lightness is predicted by the following equations:

$$L_{N_{VAC}}^* = L^* + [-0.1340 q(\theta) + 0.0872 K_{Br}] s_{uv} L^* \quad (2.19)$$

$$L_{N_{VCC}}^* = L^* + [-0.8660 q(\theta) + 0.0872 K_{Br}] s_{uv} L^* \quad (2.20)$$

s_{uv} is the chromatic saturation in terms of u', v' which predicts the effect’s strength according to colourfulness. The quadrant metric $q(\theta)$ predicts the change of the H-K effect for varying hues, and constant K_{Br} expresses the H-K effect’s dependance on

¹For readability, we have used the notation from [Nayatani, 1998].

the adapting luminance L_a . In CIELUV colour space u', v' are CIE 1976 chromaticity of test stimulus and u'_c, v'_c are chromaticities of the reference white and L_a adapting luminance, set by default to 20 as suggested by Nayatani. The remaining values are defined below [Nayatani, 1997; 1998].

$$K_{Br} = 0.2717 \frac{6.469 + 6.362L_a^{0.4495}}{6.469 + L_a^{0.4495}} \quad (2.21)$$

$$s_{uv} = 13[(u' - u'_c)^2 + (v' - v'_c)^2]^{\frac{1}{2}}, \quad \theta = \tan^{-1} \frac{v' - v'_c}{u' - u'_c} \quad (2.22)$$

$$\begin{aligned} q(\theta) = & -0.01585 - 0.03017\cos\theta - 0.04556\cos2\theta \\ & -0.02667\cos3\theta - 0.00295\cos4\theta + 0.14592\sin\theta \\ & +0.05084\sin2\theta - 0.01900\sin3\theta - 0.00764\sin4\theta \end{aligned}$$

So chromatic lightness can predict the appearance of colour, and colour spaces separate luminance from chromatic information. These two facts are important for measuring and adjusting contrast in colour images, and are used in each novel algorithm presented in this thesis. Colour properties are especially important for the greyscale conversion algorithm presented in Chapter 5.

2.4 Contrast Enhancement

When presented with an image, the task is to understand what is depicted. Through visual processing we recognize objects, interpret their reflectance properties, and resolve the scene's spatial organization. The task is made more difficult if the image is blurry or appears to have low contrast, either on account of poor eyesight or poor image quality. In the scientific and medical imaging fields, low-quality images like those arising from X-ray technology, require enhancement to increase the visibility of important areas. There are thus several methods for enhancing all types of images, here however, focus is placed primarily on the enhancement of photographic or complex images.

Contrast enhancement improves image efficacy by easing the separation of objects in a scene, simplifying the understanding of complex and self-occluding objects, by placing emphasis on salient regions and improving the impression of depth. For an in-depth review of image processing for all types of images, the reader is referred to [Gonzalez and Woods, 2002]. What follows is a review of contrast enhancement techniques, with special attention paid to local contrast adjustments and unsharp masking.

2.4.1 Global Enhancement

The contrast of an image can be interpreted through its intensity histogram. A low contrast image has a narrow histogram (Figure 2.13(a)), whereas higher contrast is associated with a histogram that covers most of the available intensities (Figures 2.13(b) & 2.13(c)). To change contrast, an image's intensity values can be re-mapped by linear scaling while adjusting to keep the average luminance constant. These so-called transfer functions, usually linear, sigmoidal or power transfer functions, are generally known as contrast stretching functions, or lightness re-mapping functions [Braun and Fairchild, 1999].

An illustration of a sigmoid transfer function and its effect on the intensity histogram is given in Figure 2.14. Histogram equalization is a statistical approach to

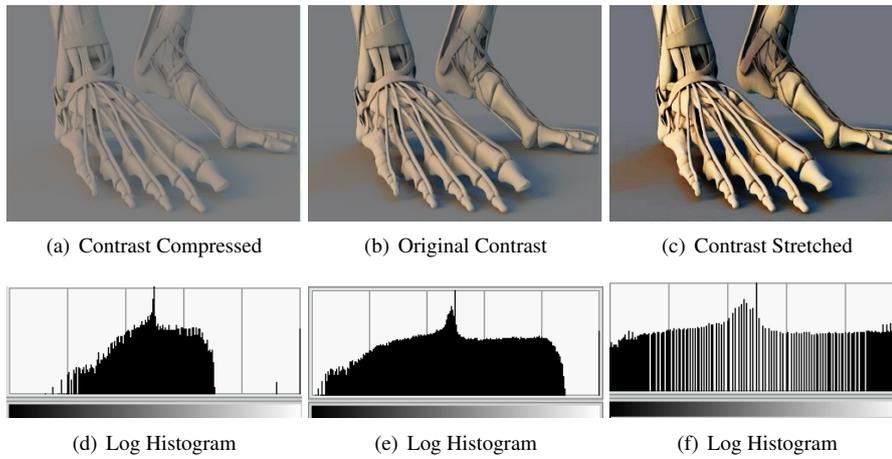


Figure 2.13: Contrast shrinking and stretching and the effect on the luminance histogram. The original image (b) has a fairly wide histogram and a nearly normal distribution (e). A low contrast image was produced by compressing the intensities to a smaller range (a) so that the histogram becomes quite narrow (d). Straight-forward histogram stretching increases the range of intensities (f) to increase contrast (c).

change the distribution of intensity values to produce a desired distribution and use the fully available intensity range. Reshaping or warping of a histogram can introduce different statistical properties that enhance contrast over low, mid and high tones. The mapping and histogram adjustment approaches work on a more global level to increase contrast of the whole image. It should also be noted that global scaling of an image does not change the contrast of its component frequencies by the same amount.

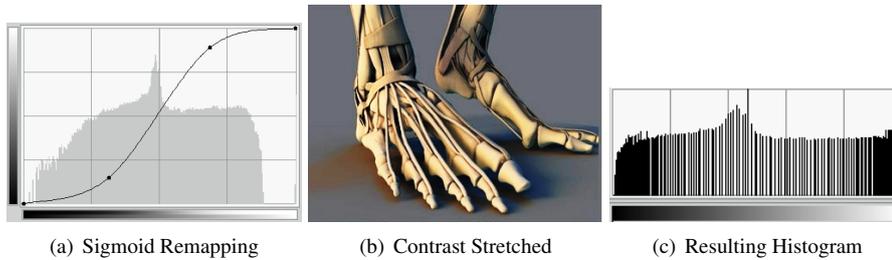


Figure 2.14: Remapping luminance with a sigmoid function to enhance contrast appearance, without interpolation.

2.4.2 Local Enhancement

Since human vision proceeds in a largely local way, we do not see an entire scene or image at once. Instead, the eye is constantly moving, perceiving local regions and using memory to put a global picture together. As such, differences between near or neighbouring regions (local contrasts) are more important to maintain than differences between all regions (global contrasts).

Local image enhancement approaches focus more on exaggerating local disconti-

nities, which makes the enhancement more structure-aware. Important differences in the image become more clear, while unimportant or uniform areas are left unaltered. Another benefit is that the histogram itself does not change globally, so enhancement is achieved without making the image appear to have a different range of intensity from the original. In addition to adapting the enhancement locally over the image, the enhancement can choose the spatial frequency of contrasts to be enhanced. One can choose to enhance very fine details, such as edges, or at lower-frequency scales, one can enhance local transitions. So, most of these algorithms set a spatial frequency size for the band of frequencies they will enhance.

The challenges of all local contrast enhancement approaches are to avoid increasing the visibility of noise - thus it is integral to differentiate between contrast due to noise, and contrast due to the scene or lighting. A large enhancement will become apparent as *visible halos* around the contrast area. This is usually controlled by working at an appropriate scale and by either human control, or the use of perceptual metrics to avoid over-shooting the addition of contrast. Lastly, colour may change in undesirable ways, so special care should be taken if the colour channels are being enhanced separately.

Since the perception of local contrast is a derivative process, the image gradients can be used to measure local contrast and can be altered to change local contrast. There are many algorithms for local contrast enhancement; in general, they measure contrast using a method similar to those described in Section 2.2, then apply computational techniques to alter the image so the contrast measure is increased or maximized.

For example, one technique is to enhance the local gradients by operating on the them directly and recovering the enhanced image by Poisson reconstruction, as is done in [Fattal *et al.*, 2002]. Note that this gradient approach is also used as a local contrast preserving tonemapping operator. Another is an optimization algorithm [Majumder and Irani, 2007] which defines an objective function that measures local contrast based on Weber's Law (Section 2.1.3) and optimizes over possible images to maximize the average local contrast.

2.4.3 Edge Sharpening Filters

Edge sharpening is a common technique in photography and Computer Graphics for the enhancement of high-frequency local contrasts. Edges are commonly extracted features in image processing, and are often manipulated to improve image clarity. The process, also known as deblurring, can increase the visibility of fine details and abrupt changes. While this does not change the contrast between the two regions per se, their relationship becomes more obvious and their boundaries more acute. The act of edge sharpening can be done by steepening the slopes of edges so they are more abrupt, or by increasing the amplitude of the edge's rate of change, as occurs in unsharp masking.

Recall the two second-order derivative operators for measuring local contrast from Section 2.2.1. Their zero-crossings detect edges, and bordering ramps describe the magnitude and polarity of the contrast (dark side positive, light side negative). This information can be used to enhance the edges in the image. The Gimp's "sharpen image" algorithm is in fact based on a Laplacian operator. It likely subtracts a fraction of the Laplacian from the original image, increasing intensity on the light side, and reducing intensity on the dark.

In the same way, the Laplacian of Gaussian (LoG) and Difference of Gaussian (DoG) can be used for contrast enhancement by subtracting either of them from the original image. However, when used for image enhancement, DoG generally means reordering the Gaussians to subtract a wider blur from a narrower one. This essentially

takes care of the subtraction, and produces an image that approximates the negated Laplacian, as was illustrated in Figure 2.9. The resulting image is then added back to the original, which effectively reproduces the subtraction of the LoG when the ratio between the two blurs is ~ 1.6 . In fact, it has been shown that when the ratio of the two Gaussians is ~ 5 , that the DoG approximates the behaviour of center-on ganglion cells in the retina, shown in Figure 2.2. It is important to note that the difference of Gaussians results in a band-limited image. In this case, containing the high-frequencies left after removing the low-frequency component resulting from the wider Gaussian blur.

2.4.4 Unsharp Masking

Unsharp masking is a simple and common technique for contrast enhancement. It was used in wet darkrooms and now is frequently used in digital image processing, print and even video to increase the perceived image quality. Unsharp masking can enhance the finest contrasts (the down side is noise enhancement) and can also enhance local contrast (when the kernel has greater spatial extent). The approach provides control over which scale of contrasts to enhance, thus sharpening edges, or enhancing broader contours.

In the darkroom, unsharp masking proceeds by first creating the “unsharp” mask. Explaining its moniker, this mask is the difference of the original image and a blurred (unsharp) version, which leaves only the high-frequency contrasts unmasked. The final contrast enhanced image is made by developing the image with normal contrast and compositing a globally contrast stretched version with the mask. In this way, the contrast-stretched areas appear only at the unmasked high-frequency regions.

As early as 1970, the ability of unsharp masking to create superior renditions was exploited, especially to improve photographs that appeared in newspapers. The main benefit is its simplicity, which meant that the technique could be “incorporated in conventional wirephoto transmitters with very little cost” [Schreiber, 1970]. The technique adapted the darkroom technique to create electronic masking on a wirephoto transmitter, and was explained as an alteration in the amplitudes of the sharp and unsharp signals.

In digital image processing, the operation is nearly identical. The “unsharp mask” is the original image minus a Gaussian blurred version, and is thus akin to the DoG method, although the first image may not be blurred. The resulting image is no longer a true mask, since it contains negative values and all the high-frequency contrast information. In fact, as shown in Section 2.4.3, the DoG, and thus unsharp masking, produces a quasi-band limited image that approximates the second spatial derivative. This image is then scaled and added back to the original to create the enhancement.

Unsharp masking contains at its core a highpass function for determining the enhancement image, also called the *contrast signal*, and another function for determining the gain values. In standard image editing software, the highpass is performed on the luminance channel or separately for each channel, by subtracting a Gaussian blurred version from the original. The sign of the highpass specifies whether there will be a lightening or darkening along an edge, creating a ‘halo’ whose size is controlled by a user specified Gaussian filter radius. The gain value is a constant user chosen amount, and an additional threshold controls minimum difference which will be enhanced.

The process for creating a contrast enhanced image $\mathcal{U}(I)$ is formulated as

$$\mathcal{U}(I) = I + \lambda h(I) \quad (2.23)$$

where λ is the gain factor that scales the highpass contrast signal $h(I)$.

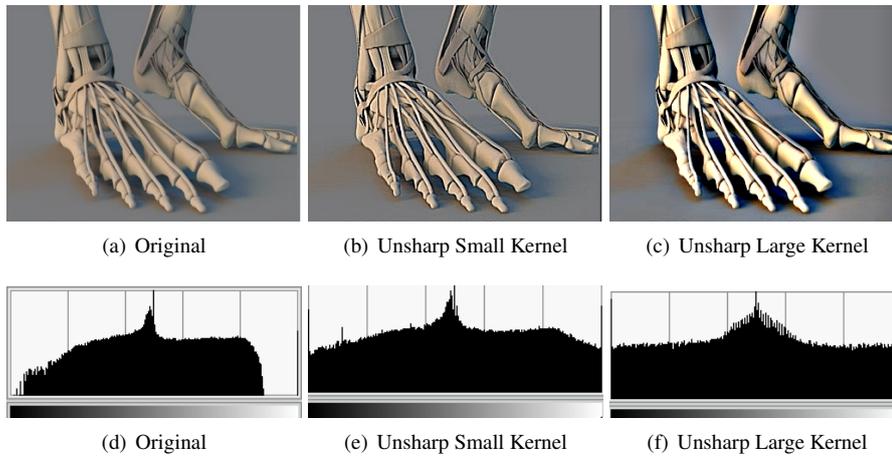


Figure 2.15: Unsharp masking to increase edge sharpness, and then to increase wider contours resulting in enhanced local contrast.

If the radius of the Gaussian blur (σ) is made larger, then wider contours are detected and enhanced. When the ramps are very wide, there is not an effect of edge sharpening, and other techniques must be done to achieve crisper edges [Leu, 2000]. Instead, enhancement of the transitional contours produces an increase in the *local contrast* between both regions, as was shown in Section 1.2.

The downside of that simple approach is that noise is amplified along with high frequencies. Research from fields like medical imaging and image processing, where input is often noisy, have developed more sophisticated and robust unsharp masking techniques. The highpass calculation has been improved from a discrete directional Laplacian to a cubic operator [Ramponi, 1998]

The most successful advancement is in calculating gain values that adapt to the image, using so-called *adaptive gain*. Mean-weighting the filter causes the medium contrasts to be emphasized more than very strong contrasts, and edge-weighting accounts for the fact that we are less sensitive to noise along the edges of an image [Ramponi, 1998]. Polesel continues in this vein by adapting to detail variance so that noise won't be overly enhanced [2000]. These adaptive gains are combined in further adaptive unsharp techniques for 2D and 3D image enhancement [Guillon *et al.*, 1998].

A study states that the quality of an unsharp techniques depends more on its adaptive gain calculation than on the highpass calculation [Badamchizadeh and Aghagolzadeh, 2004]. Another important factor is the scale of the high-pass calculation. As seen in Section 2.2.2, multi-scale image processing is motivated by visual system's treatment of different spatial frequency channels. It is common to adapt image enhancement approaches to scale-space to mimic the HVS and produce more controllable results, so it is not surprising that there are several existing multiscale unsharp masking techniques [Nowak and Baraniuk, 1998; Wang *et al.*, 2001; Kim and Allebach, 2005]

Similar unsharp masking approaches have already proven useful for colour image processing. Within their framework for extending grayscale image processing techniques to color images, Strickland *et al.* [1987] sharpen colour images by adding high-passed versions of both luminance and saturation channels to the original luminance channel. In [1997], Thomas *et al.* improve this technique by limiting artifacts due to noise and sign conflicts between the two enhancing signals according to a local mea-

sure of saliency.

Specialized unsharp masking approaches exist, and serve as inspiration to the novel approaches presented in Chapters 4, 5 and 6. Luft et al. calculate the high-pass contrast signal by applying the DoG to depth information to increase local contrast at depth discontinuities, instead of at intensity discontinuities [2006]. Finally, a type of 3D video unsharp masking is presented [Scognamiglio *et al.*, 1999] to reduce temporal artifacts caused by noise. Their enhancement adapts to the video by using a temporal filter so that edges to be sharpened have some amount of persistence over time.

2.5 Preferred Contrast and Colour Saturation

The effect of image enhancement has been explained through visual mechanisms such as lateral inhibition or the CSF transfer functions [Peli, 2000]. While the reasons contrast enhancement are effective are varied, their ability to increase perceived contrast and image quality, or at least image preference, are clear. Various studies have concluded that humans prefer an image that appears to have higher contrast between regions, making it easier to discern separate surfaces. Higher contrast may be preferred because such images look more like real scenes [Yoshida *et al.*, 2005], and possibly because the higher contrast eases comprehension of visual stimulus.

In [2003b], Calabria and Fairchild examine image preference with respect to the following three contrast manipulations:

lightness-contrast in which each test image's luminance is remapped to stretch over all contrast using a variety of remapping functions, including many of those described in Section 2.4.1.

sharpness-contrast in which the test image in *RGB* is unsharp masked using a simple version the algorithm from Section 2.4.4 with radius $\sigma = 2$.

chroma-contrast in which the chroma channel C^* of the test image in CIE $L^*C^*h^*$ is linearly scaled, and L^* and h^* preserved (see Section 2.3.2 for a colour space refresher).

The authors find that perceived contrast generally increases monotonically as contrast increases, with the exception that people find an achromatic image to have higher contrast than an image with less than 20% chroma. Image preference increases with each of the perceived contrast increases, but then decreases when the enhancement is deemed too strong. The authors combine the results of their experiments to relate perceived contrast and image preference, which connects image contrast to perceived contrast levels, meaning that image preference can be described as a function of the three contrast manipulations. So, increasing perceived contrast up to an upper threshold, increases image preference [Calabria and Fairchild, 2003b; 2003a].

Similar to the findings on sharpness-contrast, another study has supported edge sharpening procedures, finding that people prefer images with contrast increases at edges. In studies on the impact of edge sharpness by [Lin *et al.*, 2006], the perceptual quality of images was tied to a "most eye-pleasing sharpness". Images where enhanced by the LoG method, described in Section 2.4.3, and it was shown that their perceptual quality was highest when edge sharpness was increased 2.6 times the just-noticeable difference.

Chroma scaling strengthens image colourfulness, a common trend in photography, and although increased colourfulness leads to somewhat unnatural images, up to a maximum colourfulness point, they are still preferred by humans [Fedorovskaya *et al.*, 1997]. This phenomena exists partly because images are usually judged without direct reference to the original scene and memory for coloured objects can be unreliable [Bartleson, 1960], so manipulated chroma often remains unnoticed while the perceived image quality is consistently ranked higher. Fedorovskaya *et al.* [1997] through systematic psychophysical studies confirm that relation between preference and colorfulness has a shape of inverted “**u**” with the maximum preference achieved for chroma increased by 10%–20% with respect to the original image.

A chromatic increase may also lead to an image lightness increase as it is known from the Helmholtz-Kohlrausch effect introduced in Section 2.3. A direct consequence of this effect is the more recent finding that perceived contrast of colour images increases monotonically with chroma even when lightness is kept constant [Calabria and Fairchild, 2003b]. An interesting observation is that when the display dynamic range is changed (e.g., using neutral density filters) more color saturation for darker displays is required to achieve a more natural image look [deRidder, 1996]. His perceptual experiments studying the relationship between perceived image quality and naturalness were performed by varying the colorfulness of natural images at various lightness levels. At each lightness level, subjects assessed perceived colorfulness, naturalness, and quality as a function of average saturation by means of direct category scaling. Quality was found to increase monotonically with average saturation. This means that in addition to increased lightness contrast and sharpness, people prefer more colourful images.

Chapter 3

Contour Brightness Illusions

This section takes a more in-depth look at the perception of local contrast and its dependence on the type of transition between contrasting regions. To begin, Section 3.1 provides an explanation of contour terminology used throughout this discussion. Brightness illusions and the characteristic of the contours that cause them are then discussed. These give added insight into how contrast appearance depends on contours.

In Section 3.2.1, focus shifts uniquely to the Cornsweet illusion and the strength of its perceptual effect. The illusion is also shown to exist in the colour domain, which provides the necessary perceptual justification for the novel approaches for restoring lost contrast through chromatic adjustments in Chapter 4. The perceived effect of the Cornsweet illusion is strengthened when it occurs in a 3D situation, as discussed in Section 3.2.3. It is this 3D perceptual effect that justifies the 3D unsharp masking approach for enhanced rendering presented in Chapter 6.

Knowing which stimuli cause brightness illusions and knowing the strength of the effects is of course very interesting, but the reason they are studied is to elucidate the workings of visual processing. These illusions can be used to empirically explain behaviours of the visual system and in turn can serve as validation to certain perceptual theories. Thus to conclude this section on contour effects, Section 3.3 reviews several different theories explaining the connection between vision and the perceptual effect of Cornsweet contours. Returning to the introductory claim that “Unsharp masking is capable of introducing Cornsweet contours, and the perceptual effect of unsharp masking can be explained by the Cornsweet illusion”, Section 3.4 unifies the contrast enhancement algorithm and the illusion. It does so by presenting empirical evidence and observing that both depend on high-frequency peaks in luminance.

3.1 Contours and Apparent Contrast

It is known that contrast perception depends on the spatial frequency of the transition, on the contrast's magnitude and its average intensity. An abrupt transition creates the appearance of an edge, which is a thin boundary where the two differing regions meet, as illustrated in Figure 3.1(a). However, luminance transitions in the real world and in images are usually gradual and instead of producing a concise edge, produce a smooth sloping gradient, whose luminance profile resembles a sigmoid-shaped curve, as illustrated in Figure 3.1(b). The width of transition specifies the contrast's spatial frequency: a sharp edge is high-frequency, whereas a more gradual transition is mid- to lower frequency.

While apparent contrast is most directly related to a spatial intensity difference, it is also impacted by the separating contour's shape and slope characteristics [Kingdom and Moulden, 1988]. If a gradual contour blends between regions, a sharp contour does the opposite by adding distinction, thus increasing our perception of contrast. It is understandable that a sharp contour can *increase* apparent contrast; more baffling and surprising is that a contour can *create* apparent contrast [Ratliff, 1971]. Such is the case in the Craik-O'Brien-Cornsweet family of contours, where an equiluminant patch appears to have two contrasting regions purely on account of the contour [O'Brien, 1958].

The most evocative of these contours is the *Cornsweet contour*, which is a sharp discontinuity created by two opposing luminance ramps that fall back to the same luminance level away from the edge. A simple Cornsweet contour and its luminance profile are presented in Figure 3.1(c). The Cornsweet contour creates the *Cornsweet illusion*, in which the two neighbouring regions to be filled with illusory brightness that gives a sense of contrast where there is in fact none [Cornsweet, 1970]. As the reader can verify, the brightness contrast is in fact illusory; it disappears when the contour is covered.

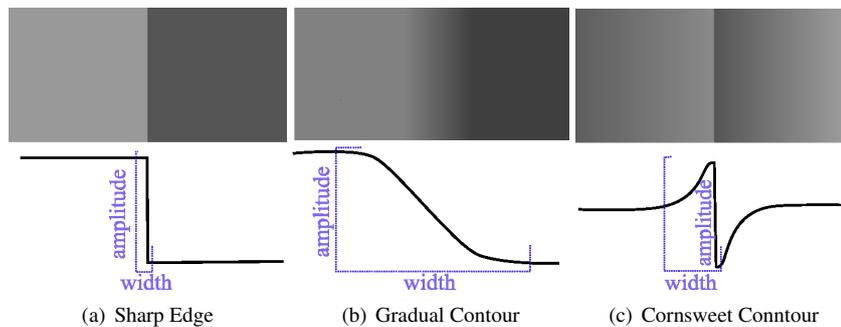


Figure 3.1: Types of local intensity change characterized by the slope of their transition and the sharpness of change at the beginning and end of the transition.

In contour brightness illusions, narrow gradient *ramps* or *peaks* in intensity make the region adjoining the light-side peak appear lighter, and the region adjoining the dark-side peak appear darker. These luminance profiles are published cases of contour affecting brightness contrast. They are studied by researchers who are trying to understand or model the luminance to brightness relationship in vision. Their existence proves that local features can increase the apparent dynamic range far beyond

that which is physically depicted.

The Cornsweet contour can be placed atop existing contrast and still affect perceived brightness. It can also exist in an asymmetric form to create the Craik O'Brien illusion, which actually was discovered before the two-sided Cornsweet effect. The luminance profiles associated with these three situations are given in Figure 3.2.

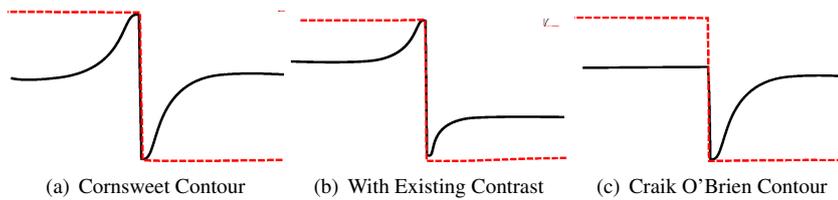


Figure 3.2: Luminance profiles (black lines) of three contours that create an unexpected apparent brightness (dashed red lines). The apparent contrast between the adjacent regions is greater than the actual contrast.

There has been research showing that the effect on perceived brightness can be spatially concatenated. This so-called repeated or cascading Cornsweet illusion is one visual stimulus made up of a sequence of similar Cornsweet contours. The perceived brightness is then a sequence of brightness steps, resulting in a sequence of illusory blocks of descending or ascending brightness [Kingdom and Moulden, 1988]. There are other contours that result in a surprising perceived brightness. For instance, the *single saw-tooth* contour replaces the parabolic ramps of the Cornsweet contour with linear ramps, for a similar illusory effect [Kingdom and Moulden, 1988].

Another related brightness illusion is the “Mach Banding” effect. However, in this case, the stimulus is a simple gradual edge and the illusion is the appearance of light and dark contours flanking the edge. The Cornsweet illusion and Mach Bands are similar phenomena. In the first, a Cornsweet contour seems to be a simple edge between two contrasting adjacent regions (Figure 3.3(a)). In the second, a gradual edge seems to have peaks of light and dark (Figure 3.3(b)).

Cornsweet Illusion The perception of illusory contrast caused by a contour with actual luminance peaks.

Mach Banding The perception of illusory contrast peaks caused by a contour with actual luminance contrast.

3.2 The Cornsweet Illusion

As stated above, the Cornsweet illusion is brought on by viewing a special type of intensity variation, which has been named the Cornsweet contour. The luminance profile of a Cornsweet contour can be generated in several ways. The center of the contour occurs at the 0 point, and the equations control the range of luminance differences and the width of the gradient ramps on either side of 0. Parameter a controls the amplitude of the contrast, which is the range between maximum and minimum luminances. The first generator uses an underlying sinusoid method whose ramp width is determined by

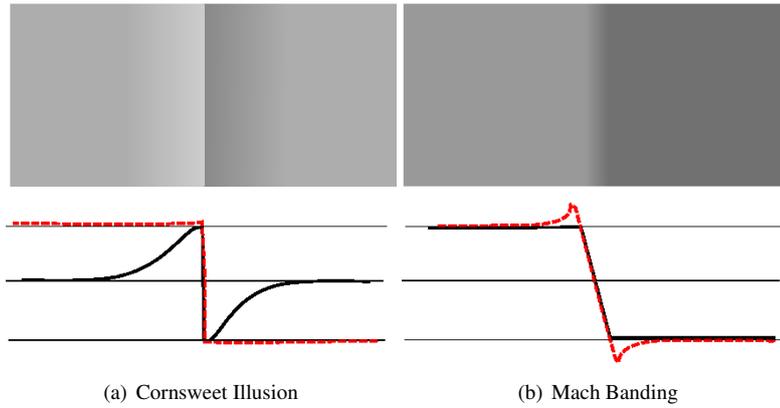


Figure 3.3: Cases where real contrast (solid black) and apparent contrast (dashed red) differ due to visual processing and human perception. The juxtaposition of two opposing gradients creates an illusion of uniform brightness differences in the Cornsweet illusion (a). A sloped edge produces the appearance of illusory bright and dark bands, named after their discoverer as “Mach bands” (b).

ω , and is defined for $-\frac{1}{4\omega} < x < \frac{1}{4\omega}$ as:

$$f(x) = \begin{cases} a \sin(2\pi\omega x) & x < 0 \\ -a \sin(2\pi\omega x) & x \geq 0 \end{cases} \quad (3.1)$$

Another method is based on a Gaussian function, so its ramp width is specified by the standard deviation σ and the Cornsweet profile is defined for $-1 < x < 1$ as:

$$f(x) = \begin{cases} a \exp(-x^2/\sigma^2) & x < 0 \\ -a \exp(-x^2/\sigma^2) & x \geq 0 \end{cases} \quad (3.2)$$

Compare the Cornsweet contour’s profile (Figure 3.2(a)) to the Laplacian of a step edge (Figure 2.8(c)) and observe that they have the same shape. This is because, as remarked by Dooley and Greenfield [1977]

“The differences between a real step and a Cornsweet edge [contour] occur principally in the low spatial frequencies; the Cornsweet edge can even be considered as a real step, from which the zero and very low-frequency components have been removed.”

Thus, by selecting the appropriate spatial parameter for the Laplacian calculation, a Cornsweet contour can be generated by taking the second-derivative of an edge, for instance using a difference of Gaussians (Figure 2.9). The amplitude of the Cornsweet contour is then proportional to the contrast of the edge. It is this approach that is used for creating Cornsweet contours in the approaches described in Chapters 4, 5 and 6.

3.2.1 Perceived Strength of the Cornsweet Illusion

Over the years since its discovery, vision scientists have performed series of quantitative experiments to analyze the perceived effect of the Cornsweet illusion. This is done

with matching experiments, where are subject attempts to adjust grey-levels to match the apparent brightnesses on either side of the contour. The effect is measured either for a vertical contour producing the percept of two adjacent squares or a radial version made of a circular contour.

Seminal work by Dooley and Greenfield [1977] resulted in the graph presented in Figure 3.4, which relates a Cornsweet contour's to the actual contrast of a step edge with the same amplitude. In these experiments, ramp width is measured in visual degrees and contrast is measured by Michelson contrast, as defined in Section 2.2. The results confirm the illusion, showing that apparent contrast approaches that which is created by an actual edge. The strength of the illusion can be seen by the amount of apparent contrast produces on account of the Cornsweet contour. The illusory effect is diminished when the ramp widths are too narrow. As amplitude increases, the strength becomes non-linear, suggesting that the illusion is stronger for lower contrasts. As expected, for an infinitely wide Cornsweet contour, the step edge contrast and apparent contrast are equal.

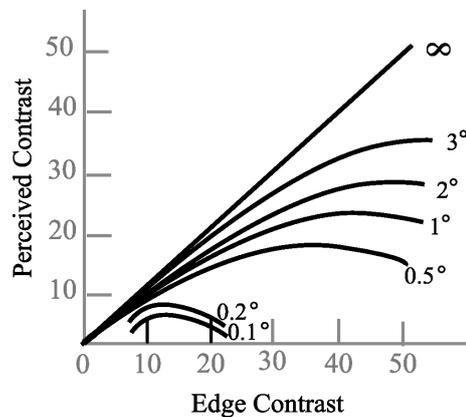


Figure 3.4: Apparent contrast of the Cornsweet illusion for Cornsweet contours whose widths (measured in visual angle) and amplitude vary. The straight line marked ∞ is equivalent to a step edge.

In the *additive cases*, the same matching experiments were performed for existing contrast in the form of a step edge, upon which a Cornsweet contour was added, as in Figure 3.2(b). The added real contrast is subtracted in the graphs shown in Figure 3.5 so that the perceived effect of the Cornsweet contour can be seen. As before, a wider ramp creates higher apparent contrast. By studying the effects of greater real contrast, [Dooley and Greenfield, 1977] concluded that the strength of the effect decreased as real contrast increased.

One difference between the two types of stimuli is that the radial version has a perceived interior (the inner disk). In the asymmetric case (the Craik-O'Brien contour), the perceived contrast is greater in the radial stimulus version when the fall-off ramp is on the outside. Additionally, empirical evidence by the author (unpublished) shows that the effect indeed occurs for other types of contours, like that of the sample matching experiment shown in Figure 3.6.

When considering spatial influence, the finding is that the strength of the illusion is consistent under changes to transition width, so long as the ratio between

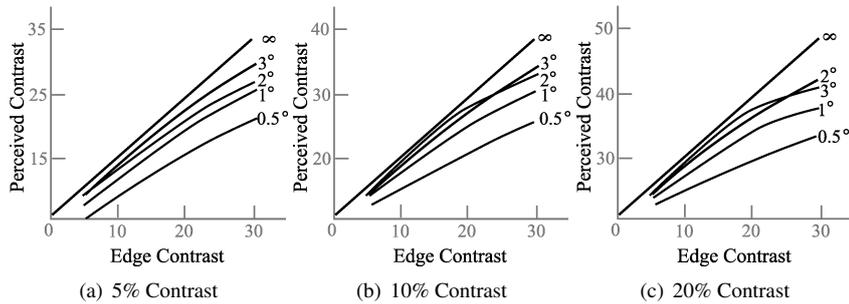


Figure 3.5: Apparent contrast of the Cornsweet illusion for *additive* Cornsweet contours whose widths (measured in visual angle) and amplitude vary. The underlying step edge adds 5, 10 or 20 percent real contrast, which is subtracted from the vertical axes for comparison purposes.

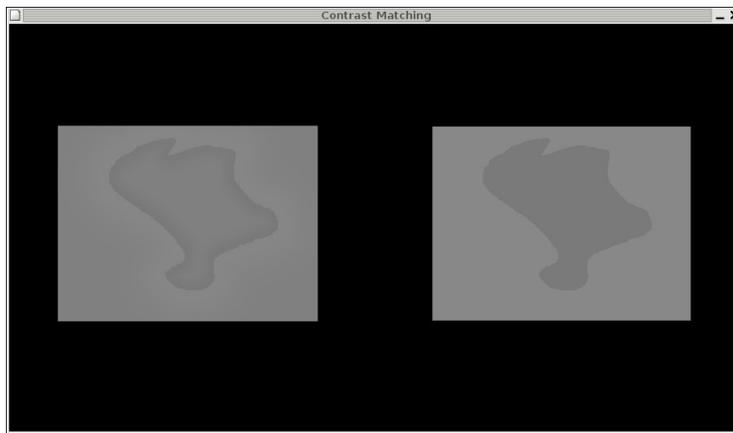


Figure 3.6: A simple matching experiment for quantifying the perceived contrast of the Cornsweet illusion. In this example, the users increase and decrease both the inner and outer luminances on the right-hand side to match the brightness they see in the left-hand image, which is created by an arbitrarily shaped Cornsweet contour.

total stimulus width and the transitional ramp widths is preserved. The temporal aspects of the illusion have also been considered, and it has been found that when the visual system has time to adjust to the contour and increase its understanding of the stimulus, that the illusory brightness fades [Wachtler and Wehrhahn, 1997; Davey *et al.*, 1998].

3.2.2 The Chromatic Cornsweet Illusion

Wachtler and Wehrhahn examined the strength of the Cornsweet illusion created by achromatic and chromatic contours [1997]. A chromatic Cornsweet contour can be created using the same signal shape as in the achromatic version by changing hue or saturation and adjusting the values to preserve luminance. In their studies, colour contrast was created by modulating yellow according to the tradition Cornsweet con-

tour profile. To illustrate the effect of chromatic Cornsweet contours, two stimuli have been created by shifting hue and shifting saturation, while trying to preserve lightness, as shown in Figure 3.7. Empirical results by the author show that chromatic Cornsweet contours can also be created by adjusting chromaticity only, and preserving lightness and hue, however, the perceived brightness illusion is slightly weaker. There is a consensus that the effect of achromatic Cornsweet contours is one order of magnitude greater than the chromatic version [Ware and Cowan, 1983; Wachtler and Wehrhahn, 1997].

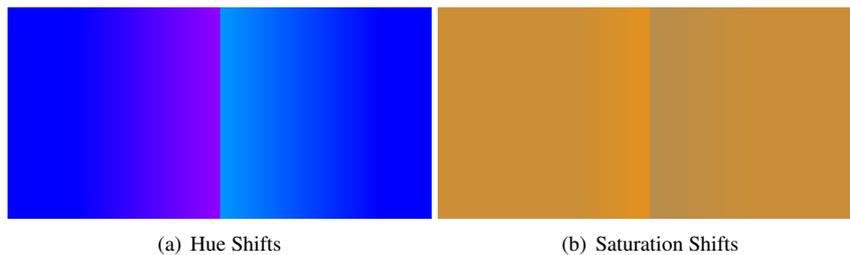


Figure 3.7: Chromatic Cornsweet contours created by hue shifting and saturation shifting. These examples are nearly isoluminant, and are here mainly for illustration.

3.2.3 The 3D Cornsweet Illusion

Purves et al. [1999] present empirical findings on the Cornsweet illusion in a 3D scene, instead of as a 2D stimulus. In a scene, transitions in intensity can arise from changes in surface reflectance/texture variation, curvature, attenuation with distance, gradients of illumination, penumbra (occlusion) and transmittance (partial occlusion). If the Cornsweet contour seems to occur for one of these reasons, it is reinforced by compatible visual cues and its effect is strengthened.

Through perceptual experiments, the Purves group compares the effects of different Cornsweet contours, including the standard vertical contour used in the studies mentioned above. Their goal is to prove that the illusion occurs because humans probabilistically reason about how the luminance profile could arise in the real world, and that the Cornsweet contour will appear to be caused by two adjacent regions of contrasting uniform lightness that occur in a natural scene. To investigate this theory, the researchers created a set of stimuli that increase the probability that the Cornsweet contour is in fact a variation in illumination or reflectance in a real world scene.

The first stimulus encourages the interpretation of a painted on contour by adding a grey frame that prevents the Cornsweet contour from reaching the border of the stimulus. Another stimulus shows the Cornsweet contour as the result of reflected illumination differences of two curved surfaces, shown in Figure 3.8(a). Two different horizontal Cornsweet contours are created: one that can be interpreted as stemming from 3D geometry shading (Figure 3.8(b)), and the other that can not (Figure 3.8(c)). Finally, the last stimulus is an image of a 3D scene where the contour occurs at the indented edge of a curved cube, and the visual cues surrounding the cube all confirm the interpretation of the Cornsweet contour as caused by shading, as shown in Figure 3.9. These stimuli are created to increase the probability that the Cornsweet contour is in fact a variation in illumination in a real world scene.

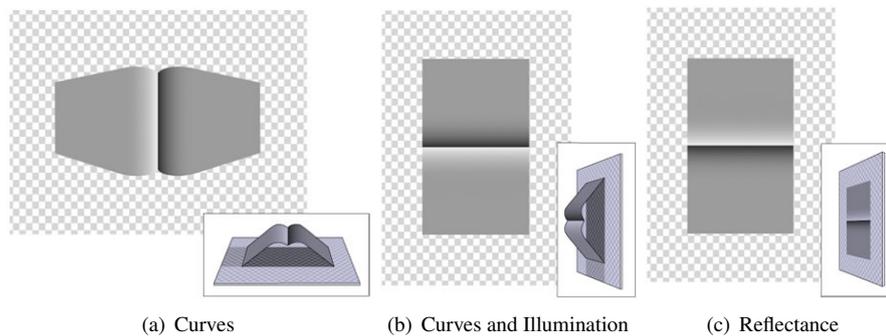


Figure 3.8: 3D Cornsweet contours and their possible real-world origins. Images from the *Purves Lab*, Dale Purves website at Duke University.

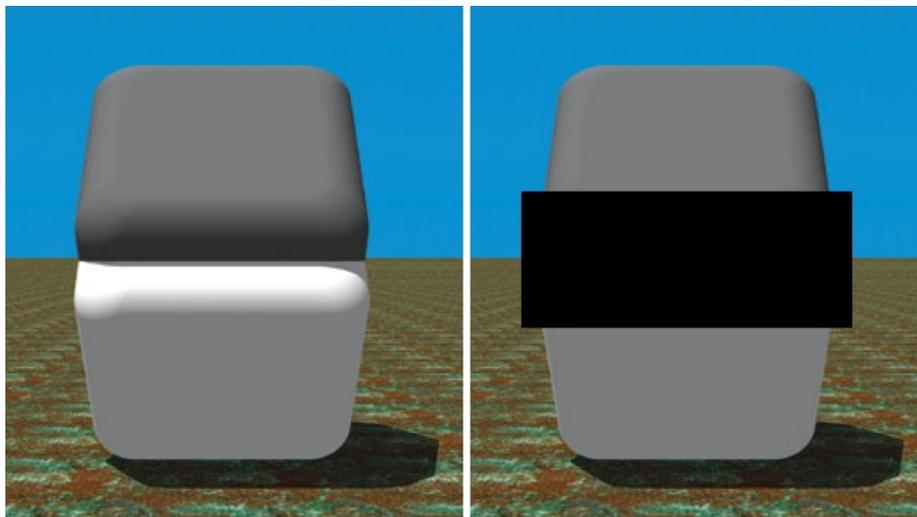


Figure 3.9: Illusions from Purves et al. [Purves *et al.*, 1999] showing that a Cornsweet contour in a 3D scene produces a much stronger perceptual contrast than simply showing the standard Cornsweet contour stimulus. The top face of the cube appears much darker than the bottom face when in fact they are nearly identical. Image from the *Purves Lab*, Dale Purves website at Duke University.

Their results show that indeed, the Cornsweet illusion is strengthened by visual cues that enforce the interpretation of the contour as an intensity variation caused by lighting or material differences. The first evidence is that the illusory effect of the painted-on contour is 68% weaker than the simple vertical contour stimulus. This result shows that the illusion is weakened if the Cornsweet contour seems to be part of a surface's texture, or if it becomes discordant with other aspects in the image.

The two horizontal stimuli show an increase of 9.5 – 19% over the strength of the standard stimulus. The stronger effect is for the stimulus shown in Figure 3.8(b), which has an easier 3D interpretation. The contour seemingly caused by two curved surfaces (Figure 3.8(a)) has a 30% stronger perceived effect on apparent contrast. This

large improvement over the standard Cornsweet contour's effect could be explained by perspective foreshortening of the adjacent regions, which reinforce a 3D interpretation. Thus it would seem that the illusion's salience is increased when it implies depth by incorporating perspective projections.

Astoundingly, and as can be clearly seen by the reader when he or she covers the Cornsweet contour in Figure 3.9, the 3D scene creates a 167% increase in apparent contrast over a standard Cornsweet contour. These findings show that to be most effective, the Cornsweet contour's gradient orientation should be consistent with lighting, as well as other cues like texture and background, and any other cue that would indicate normal viewing conditions. They also imply that if a Cornsweet contour is added to an image for enhancement purposes, it must take care not to appear superficial or at odds with the interpretation of the scene.

3.3 Perceptual Explanations

The purpose of studying perceptual illusions, other than for image enhancement of course, is to test different theories on visual processing. These brightness illusions can help to confirm or disprove theories on how humans see, since they are cases where the appearance is unlike what is usually expected.

Dooley and Greenfield [1977] state that a spatial vision model would predict that the eye ignores the missing low-frequencies in the Cornsweet contour. By ignoring them, the human interprets the high frequency component just as he would interpret a step edge. However, this theory does not account for the diminished strength of the illusion in additive cases, when underlying real contrast is present.

[Wachtler and Wehrhahn, 1997] suggest that the Cornsweet illusion is caused by humans' lack of sensitivity to gradients at high-frequencies. This theory is based on the Contrast Sensitivity Function (CSF, Section 2.1.3). It is postulated that contour illusions occur because ramps are not distinguished from the uniform regions, and are thus considered uniform as well.

The lightness integration theory suggests that the ramps of the Cornsweet contour are not detected because of the low-level processing of the center-surround cells described in Section 2.1.2. Due to the behaviour of these cells, the spatial aspects of visual processing are highly sensitive to abrupt changes, and not so sensitive to gradual intensity differences, especially those that occur over narrow spatial ranges. Thus computational models like the Retinex model, or even the Difference of Gaussians, are capable of predicting the apparent contrast of Cornsweet contours.

The filling-in hypothesis postulates that local contrasts are processed in the cortex and between them is a filling-in or spreading as the input information fills the empty areas adjacent to the local contrast. The theory is temporal since it takes time for the illusory brightness to propagate to the adjacent regions. Thus, if modeled by filling-in, then the decrease in the illusion's strength over time may be accounted for [Davey *et al.*, 1998]. However, other research that studied the Cornsweet illusion for different spatial frequencies found that the temporal aspects of the Cornsweet illusion are not consistent with the filling-in theory [Devinck *et al.*, 2007]. These researchers found that the illusion is more consistent with the theory of spatial filtering in the human visual system.

Finally, Purves *et al.* [1999] provide a holistic explanation of various brightness illusions by arguing that brightness is perceived probabilistically according to the most likely scene that causes a visual stimulus and humans' prior experience. In this theory,

the brightness of a Cornsweet contour is perceived by predicting the probable combination of geometry and illumination that would produce the gradients. If additional visual cues imply that the contour is a reflected texture on an equiluminant flat surface, then the apparent contrast is reduced. On the other hand, if visual cues imply the contour is a product of reflectance differences, or different illumination, then the apparent contrast is much higher.

3.4 Unsharp Masking and the Cornsweet Illusion

Now that all of the background has been presented, it is time to return to the second principle of this thesis.

Second Principle Unsharp masking is capable of introducing Cornsweet contours, and the perceptual effect of unsharp masking can be explained by the Cornsweet illusion.

Section 2.4.4 describes unsharp masking as an image processing technique that proceeds by approximating the second-order derivative of an image's intensity, which is then introduced into the original image to produce a local contrast enhancement. Recall that Section 3.2 shows how a Cornsweet contour that causes the perception of a Cornsweet illusion can be created by taking the second-order derivative of a step edge, illustrated in Figure 2.8. So this confirms the Second Principle, but only for step edges. What about other types of local contrast? These contrasts can be sharp or blurred (describing the ramps on either side of the intensity change), and abrupt or gradual (describing the transition width of the intensity change).

Consider the case of an abrupt but slightly blurry local contrast, whose profile would look like the left-most curve in Figure 3.10(a). When unsharp masking is applied, the second-order derivative (the middle curve) is also abrupt, yet its peaks are slightly rounded. The enhanced contrast (right-most curve) has now acquired the intensity peaks of the second-order derivative. Similarly, consider the unsharp masking of a gradual local contrast that sharply ends, like the left-most curve in Figure 3.10(b). In this case, the second-order derivative (the middle curve) is not abrupt but is gradual, and the peaks are sharp and separated by enough distance so they occur where the contrast starts and finishes. The unsharp masked version (right-most) also acquires peaks at the edges of the contrast.

The argument made here is that these second-order derivatives are types of Cornsweet contours that have different sharpness and spatial extent. They do not create as strong an effect on apparent contrast as the standard Cornsweet contour, whose profile is shown in Figure 3.2(b), but their shape is similar nonetheless, and it follows that they will be perceived in a similar way. By adding these varied Cornsweet contours back to an image, the original local contrasts acquire small intensity peaks, making them produce weak Cornsweet illusions. Or more accurately, the peaks create small illusory changes to the brightness of the two adjacent regions, which makes the local contrast appear greater than it actually is.

Following this line of reasoning, the same aspect of visual processing is responsible for both the Cornsweet illusion *and* for making unsharp masked images appear more contrasted. This concept is illustrated in Figure 3.10 using dotted red lines to represent apparent contrast. It is possible too, that in the case shown in Figure 3.10(b), that the contrast is formed by two spatially separated Craik O'Brien contours (Figure 3.2(c)), instead of by a sloping Cornsweet contour. Whichever the case, the apparent contrast

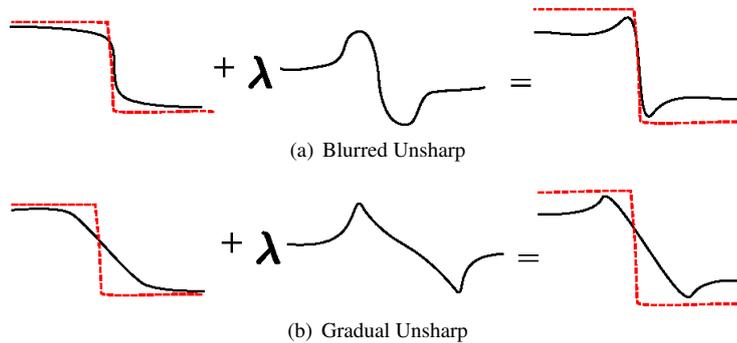


Figure 3.10: Depiction of unsharp masking on a blurred edge (a) and a wider gradual edge (b). The Laplacians define Cornsweet contours whose shape varies from the standard Cornsweet contour, but the characteristic intensity peaks are preserved. For a blurred edge, the peaks are smoother, and for the gradual edge, the peaks are spatially distant. Still, their addition to the original edge create intensity peaks that probably increase apparent local contrast in the same way the standard Cornsweet contour does, albeit with a weakened effect. For a gradual edge, the process may be more akin to adding two Craik-O'Brien contours.

produced is equivalent. To further illustrate this point, unsharp masking has been applied to each of these three types of edges, and it is clear that changing the intensity profiles of the edges to include intensity peaks does indeed increase the apparent local contrast by making the adjacent regions appear lighter or darker, Figure 3.11.

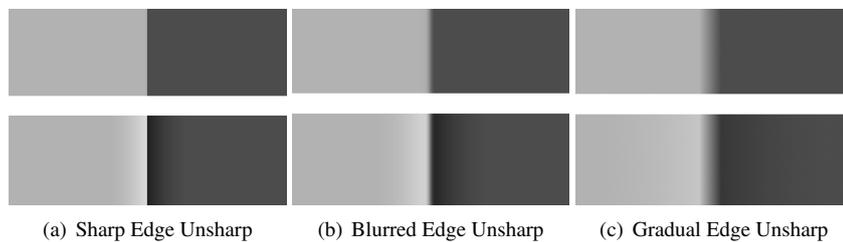


Figure 3.11: Application of unsharp masking on a step edge (a), a blurred step edge (b) and a wider gradual edge (c). Notice that Mach Bands appear at the sides of the gradual edge, but that unsharp masking nonetheless changes the apparent brightness of the dark and light rectangles (c).

A naïve explanation of unsharp masking would be that the increase in high frequencies makes the edge look sharper. Unsharp masking may be able to sharpen a blurry step edge, but it can not sharpen a gradual edge whose true sharpening would require direct steepening of its slope. Photographers understand this difference, and in digital photography, unsharp masking with a wide kernel is done to increase *local contrast*, not to increase *edge sharpness*.

Some have explained unsharp masking as the addition of visible Mach Bands to either side of local contrasts. The theory is that the perceived Mach Bands become even more pronounced atop of the true intensity bands. Indeed, the process of adding

intensity peaks near an edge may be inspired by the desire to depict contrast as it is seen, in this case the illusory Mach Bands. However, it is not accurate to think that Mach Bands can be “added” to an image, since they are purely perceptual entities and don’t themselves exist physically. In addition, when proper scale values are used, unsharp masking enhances contrast without creating visible bands or contours. Furthermore, unsharp masking is most effective when the added contrast peaks are imperceptible, which is yet another characteristic of Cornsweet contours and the Cornsweet illusion.

These claims that unsharp masking and the Cornsweet illusion share a perceptual explanation are based on observation of similar characteristics and empirical evidence. It is certainly undeniable that unsharp masking an edge gives it the similar high-frequency properties that characterize the illusion-causing Cornsweet contours. In the context of image enhancement, Chapter 5 provides compelling evidence that the two are related by showing how unsharp masking can create illusory contrast between isoluminant areas in greyscale images. Even more so, Chapter 6 performs unsharp masking of a 3D scene to reproduce the 3D Cornsweet illusion shown in Figure 3.9.

Chapter 4

Restoring Apparent Contrast To Tone Mapped Images

High Dynamic Range (HDR) images accurately describe the wide range of luminance visible in the real world. Because their dynamic range is broad enough to represent the true range of luminosity in a scene (between 3 to 12 orders of magnitude), HDR images capture details that are perceived by the human visual system (HVS) but missed by standard photographic techniques.

HDR images are well known to the computer graphics research community, and the recent introduction of HDR image creation and editing capabilities into most common image editing software ensures that HDR images will become an increasingly common form for storing and manipulating visual information. Additionally, the high quantity of information in HDR images can prevent editing artifacts and should improve the performance of image processing algorithms. As such, there is new interest in HDR processing techniques and methods for exploiting the expanded information contained in HDR images.

The problem of mapping the visible range of intensities to a much more restricted range is not new or untried. Painters endeavor to artistically map what they perceive to what can be depicted by the limited light range achievable with canvas and paints [Durand *et al.*, 2002]. Some employ colour concepts to create the impression of a wider dynamic range and greater contrast, as shown in Section 4.1. Traditional photographers encounter this problem both when taking a photograph and when printing negatives to lower dynamic range photographic paper. Inspired by photographic practices, Reinhard *et al.* introduce a tone mapping operator that extends the printing techniques of renowned photographer Ansel Adams [Reinhard *et al.*, 2002]. Section 4.2 outlines various different tone mapping operators and their characteristic strengths and weaknesses.

In novel work that adds a chromatic enhancement step after tone mapping, the LDR depictions of HDR images appear to have greater overall contrast and more evident details. Two techniques are given: the first uses distortion metrics to measure the contrast loss between the HDR and LDR images (Section 4.3.2) and the second provides an automatic approach for adding chromatic Cornsweet contours with a modified unsharp masking algorithm (Section 4.3.3).

4.1 In Art: Overcoming a Limited Dynamic Range

The task of mapping the wide range of real world intensities to a more restrictive medium is challenging, yet painters are expertly able to create convincing and detailed depictions of real world scenes. Through their collective understanding of human vision and perception, they have developed various artistic techniques to reproduce the perceptual effect of viewing a true scene, including the impression of a vast illumination range [Livingstone, 2002].

Although realism was the historical goal giving rise to such artistic techniques, realistic paintings are not necessarily photo-real. In painting and photography, intentional chroma operations are used to affect the appearance of global contrast and detail contrast. For instance, creating a controlled halo at large feature boundaries increases the perceived brightness difference between the feature and its surround, helping the HVS perform the cognitive task of segmenting features from the background [Tumblin, 1999].

The French colourists of the 19th century were inspired by strong lighting changes and bold uses of colour to emphasize details and adornments. In Floyd Ratliff's essay on color and painting [Ratliff, 1996] an example is given by Eugène Delacroix's *Femmes d'Alger dans leur appartement*, a photograph of which is shown in Figure 4.1. In this painting, the following colour techniques are displayed: the juxtaposition of carefully chosen hues to enhance contrast, the use of coloured shadows and the elimination of earth colors from the palette (unsaturated colours).



Figure 4.1: E. Delacroix makes a bold use of colour to emphasize details and adornments in his 1834 painting *Femmes d'Alger dans leur appartement*. Original at Musée du Louvre Paris.

Heightened contrasts can be achieved by contrasting hues or contrasting colour saturation. By shifting shadows to blue, and then surrounding them with yellow, painter Claude Monet juxtaposed opponent colours to increase apparent contrast, as shown in

Haystack, Morning Snow Effect (Meule, Effet de Neige, le Matin) pictured here in Figure 4.2(a). He also typically used chroma to distinguish nearly shapeless flower details that are isoluminant with their leafy background to create an effect of movement, shown here in Figure 4.2(b). The perceptual effect of his paintings and his use of colour are elucidated in “Vision and Art: The biology of seeing” [2002].

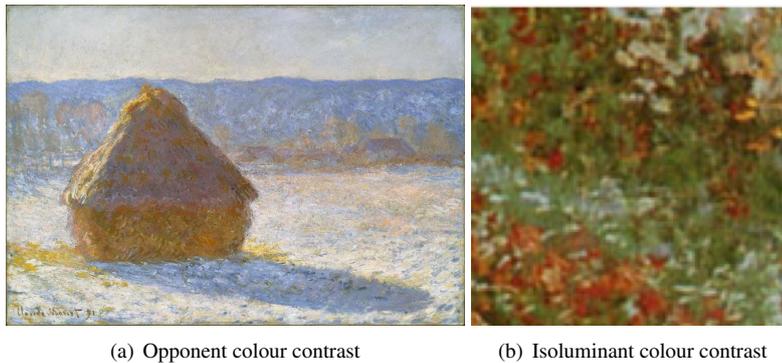


Figure 4.2: Monet’s use of colour: (left) chromatic contours of juxtaposing blue and yellow are placed around shadows to increase apparent contrast in his 1891 painting *Haystack, Morning Snow Effect (Meule, Effet de Neige, le Matin)* from the Museum of Fine Arts, Boston (a). In this detail from his 1873 painting *The Artist’s House at Argenteuil* (Art Institute of Chicago), the flower details are created almost entirely through chromatic variations over a mostly isoluminant area.

4.2 Tone Mapping Background

Tone mapping is the first and most developed research area in HDR image processing [Reinhard *et al.*, 2005]. Tone mapping compresses the wide dynamic range to a narrower range for display and aesthetic purposes thus creating an LDR depiction of an HDR image. For a majority of existing tone mapping operators this is achieved through the reduction of physical contrast in LDR images. In general tone mapping operators can be divided into global image processing tools, algorithms based on psychophysical models and methods derived from perception theories. For a full description of HDR images and various tone mapping techniques, the reader is advised to consult [Reinhard *et al.*, 2005].

The histogram adjustment method presented by Ward *et al.* is one global algorithm that compresses image areas with few samples while preserving areas that are more varied [Ward Larson *et al.*, 1997]. This method performs a compression of luminance range, however can result in the loss of fine details.

The psychophysical approach is either based on modelling the sigmoid response of rods and cones [Pattanaik *et al.*, 2000; Reinhard and Devlin, 2005] or is based on the power-law [Stevens and Stevens, 1960] relationship between brightness and underlying luminance [Tumblin and Rushmeier, 1993]. The main objective is to match the brightness of a scene perceived on a display and its real counterpart for any lighting condition. Often a significant compression of the luminance range leads to a visible loss of fine details. To prevent the luminance range compression from lowering the per-

ceived global contrast, a tone mapping operator by [Spencer *et al.*, 1995] reintroduced glare as a perceptual cue of extremely high luminance.

Further research led to the development of more sophisticated algorithms that preserve local details. Methods based on contrast theories [Jobson *et al.*, 1997; Fattal *et al.*, 2002; Mantiuk *et al.*, 2006b] operating on luminance gradients perform particularly well. Of particular interest are tone mappers that involve intrinsic image models [Barrow and Tenenbaum, 1978; Arend, 1994]. These methods separate the HDR image into illumination and reflectance layers. The high contrast of the illumination layer is usually reduced by scaling, while the details layer (assumed to be of low contrast) is preserved. While the idea is simple, the separation into intrinsic image layers is not trivial. Tumblin *et al.* introduced such an approach and assume explicitly provided layers, as is the case for synthetic images [Tumblin *et al.*, 1999]. This work was followed by several methods for automatic layer separation. The LCIS operator [Tumblin and Turk, 1999] separates the image into large scale features (presumably illumination) and fine details. Following the LCIS technique, the bilateral filter method presented by Dorsey and Durand achieves a successful separation method into “base” and “detail” layers, the first of which can then be scaled to compress the luminance range [2002].

Each tone mapping operator takes form as a collection of certain image processing operations, whose impact on the perceived image quality or fidelity to the real world appearance is not well understood. Recent psychophysical studies attempt to evaluate tone mapping operators in terms of subject preference or fidelity of the real world scene depiction [Kuang *et al.*, 2004; Ledda *et al.*, 2005; Yoshida *et al.*, 2005]. In such studies each operator is treated as a black box and its performance is compared on the whole with respect to other operators, without an attempt at understanding the reasons for subjects’ judgments. While some studies of tone mapping operators go further and take into account the reproduction of overall brightness, global contrast or details in dark and bright image regions [Ledda *et al.*, 2005; Yoshida *et al.*, 2005], they remain focused on comparing the operator performance for each of these tasks.

4.3 Beyond Tone Mapping

A tone mapping operator is often designed to produce images that look pleasing or to obtain a perceptual match between the image and the corresponding real world scenes. The success of meeting these goals depends heavily on particular HDR image characteristics and as such, it is difficult to single out one existing operator that consistently performs best [Reinhard *et al.*, 2005]. A user must choose the operator that he or she prefers for a particular image, and adjust tone mapping parameters according to the image content and anticipated display medium.

But how has the HDR image have been transformed by tone mapping and in what way has the perceived image contrasts changed? The spatial frequency theory of vision states that humans process a base layer of global contrasts and one or more layers of higher-frequency detail contrasts, as discussed in Section 2.1.3. The problem is that both global contrast and detail contrast have been altered. All tone mapping balances a trade-off between preserving global contrast and preserving details. When comparing the original and mapped images, one can measure distortions at both global (base) and detail levels. Global contrast is closely related to image comprehension, which according to Gestalt theorists, involves the cognitive task of separating the image into recognizable objects, most importantly, the separation of foreground objects from the background [Livingstone, 2002]. As such, a decrease in global contrast may make

comprehension of the LDR image more difficult, indicating a loss in visual communication efficacy. Detail distortions could occur when high-frequencies are under- or over-preserved. However, it is much more common to notice a loss of details.

The goal of the following two novel approaches is to restore the impression of base and detail contrast, while preserving the tone mapped luminance. Changes are made in the LDR image after tonemapping to create an enhanced, but still realistic depiction of the HDR image. Given an HDR and tone mapped image pair, locations of contrast loss are measured and chromatic Cornsweet contours are added accordingly to restore original contrast information. Cornsweet contours created by chroma shifts have the ability to increase perceived contrast, as shown in Section 3.2.2. Creating such contours in the LDR image by chroma scaling can augment detail visibility, and encourage the prominence of foreground objects, thus reclaiming the loss in perceived global contrast. This approach is motivated by [Tumblin and Turk, 1999], who asked how “a more thoughtful treatment of colour could better exploit the boundary and detail information”.

Section 4.3.2 describes the *distortion method* as originally published. A deepened understanding of Cornsweet contours and their addition by unsharp masking led to a series of simplifications. As described in Section 4.3.3, the newer *unsharp method* is single-step and automatic that creates higher quality results. Both techniques involve chromatic scaling to create local contrast, as explained in Section 4.3.1.

4.3.1 Colour Contrast and Chromatic Scaling

Given that colour is an inherent attribute of image quality, one would assume that tone mapping operators perform some colour enhancement. However, the majority of tone mappings compress only luminance values, and are not concerned with color issues. Two notable exceptions are the iCAM model [Fairchild and Johnson, 2003] and the multi-scale adaptation model [Pattanaik *et al.*, 1998], both of which are advanced image appearance models that incorporate colour appearance modeling [Hunt, 1995].

The image enhancement aspect of this work relies extensively on the use of colour in imagery and is related to image recolourization. Colour is a prominent attribute for effective visual communication and its use is addressed in a variety of fields including colour appearance modelling, scientific visualization and image processing. The work most closely related to adding chromatic contours is image recolouring, which transfers colours between images, introduces colours into a greyscale image or quantizes the number of colours in an image [Reinhard *et al.*, 2001; Rasche *et al.*, 2005].

Colourfulness Attribute of a visual sensation according to which the perceived colour of an area appears to be more or less chromatic [Fairchild, 2005].

Chroma Colourfulness of an area judged as a proportion of the brightness of a similarly illuminated area that appears to be white or highly transmitting [Fairchild, 2005].

Saturation Colourfulness of an area judged in proportion to its brightness [Fairchild, 2005].

Technically, chroma has a clear mathematical definition, and can be easily manipulated by processing images in certain perceptual colour spaces, like CIE $L^*U^*V^*$ or L^*C^*H . For the reader who requires a background, Section 2.3.2 introduces colour spaces and terms. Chroma scaling strengthens image colourfulness, and is a common

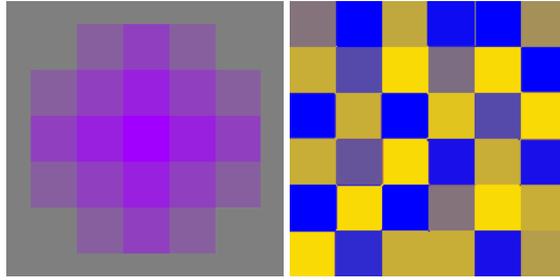


Figure 4.3: *Contrast of saturation and contrast of complements with varying luminance* (Weber State University).

trend in photography. As surveyed in Section 2.5, human preference is for images with strong colorfulness, which may even lead to their somewhat unnatural but still preferred look.

Colour contrast is the perceived difference that arises from the juxtaposition of two different colours. Such contrast is ideal for restoring lost luminance contrast resulting from HDR compression because colour is a flexible and aesthetic property, and because colour is often treated too casually by tone mapping operators. An additional reason for working with colour is to limit changes to the high quality luminance compression resulting from the tone mapping.

Chromatic Cornsweet contours arise from contrasting saturation which is adjusted by scaling chroma. Positive scaling increases colourfulness, while negative scale values move the colour towards a saturated opponent colour. Of the seven general types of colour contrast [Itten, 1961], chromatic scaling creates *contrast of saturation* and enhances *contrast of complements*, illustrated in Figure 4.3.

Images are converted to the approximately uniform perceptual CIE $L^*u^*v^*$ colour space, where axis L^* represents perceived lightness and u^* and v^* is roughly decorrelated chromatic axes coinciding with red/green and yellow/blue opponent hue pairs, as defined in Section 2.3.2. These spaces are ideal for image enhancement because luminance is related to a perceptual scale of lightness and because the space provides a correlate of chroma, defined as

$$C_{uv}^* = \sqrt{u^{*2} + v^{*2}} \quad (4.1)$$

which can be interpreted as a perceptual measure of colourfulness with respect to a white of similar brightness [Hunt, 1995].

Chroma is then scaled to increase or decrease the perceived colourfulness of the colour without changing hue angle h_{uv} or lightness L^* , because the scale value cancels out leaving the hue angle unmodified (although in extreme cases, the angle may flip to the opponent colour), and L^* is unchanged. Before performing these alterations, the image must have high quality chromatic information, otherwise artifacts may be introduced. Any undefined colour information in the tone mapped LDR image I , pixels with undefined chroma or chroma that is drastically different from the HDR image, is reset to properly scaled HDR u^* and v^* values, thus reintroducing chromatic information that may have been lost due to clamping. A minimum reliable chroma value is set so as not to scale any unreliable pixels, thus avoiding the enhancement of noise. Additionally, the white point of the image drastically impacts the overall impression of the image, so pixels within the 99th luminance percentile are not modified.

4.3.2 Distortion Method

Previous work by Smith et al. [Smith *et al.*, 2006] presented the first feature-based characterization and objective perceptual metrics for *Global Contrast Change* and *Detail Visibility Change* between an HDR image and its tone mapped LDR counterpart. Global contrast change is measured as a characteristic defined by the shape of the tone mapping function, thus removing the emphasis on extreme brights and darks which have less impact on the impression of global contrast. Detail Visibility Change as the reduction, disappearance or exaggeration of high frequency contrasts in the LDR image compared to the HDR original. A number of other perception-based visible difference (fidelity) metrics for image pairs have been developed, mostly for image compression and color reproduction applications. Refer to [Winkler, 2005] for a recent survey of such metrics.

The two distortion metrics provide an indicative characterization of these operators in terms of global contrast and detail preservation in dark and light regions. Global contrast is denoted by a single value C , and detail visibility is a per-pixel mask $\bar{T}_{i,j}^*$ of just noticeable differences. For details on these particular calculations please see [Smith *et al.*, 2006].

Using these measures, the chromatic component of the LDR image is scaled to compensate for the distortions, and the enhanced image I' is defined as

$$I' = (L_{uv}^*(I), mC_{uv}^*(I), h_{uv}(I)) \quad (4.2)$$

for scale values $m = m_G$ or $m = m_D$. When $m \geq 1$, chroma increases (colour becomes saturated with respect L^*), and when $m < 1$, colours become desaturated until they are achromatic and then become saturated in the opponent hue. The alteration is done by scaling both u^* and v^* by m , because

$$mC_{uv}^* = \sqrt{(mu^*)^2 + (mv^*)^2} \quad (4.3)$$

Global Contrast Restoration is inspired by a photographic and artistic technique named “countershading” which encourages image efficacy and creates the impression of larger global contrast. Countershading is the juxtaposition of opposing gradients to create an exaggerated difference at a feature boundary, often the boundary between foreground objects and the background. With its opposing gradients, countershading can be interpreted as the addition of a wide Cornsweet contour over large-scale image features. It is a technique used by renowned photographer Pete Turner, who characteristically creates photographs with saturation gradients applied to backgrounds, as shown in Figure 4.4(a).

To create a countershade effect in the tone mapped image, a chromatic Cornsweet contour is added along the border between the foreground and background of the image by generating chroma scale values m_G defined as:

$$m_{Gi,j} = \begin{cases} a \cdot \exp(-d^2/\sigma^2) + 1, & I_{i,j} \in \text{Segment A} \\ a \cdot (1 - \exp(-d^2/\sigma^2)) + 1, & I_{i,j} \in \text{Segment B} \end{cases} \quad (4.4)$$

where d is the shortest distance to the foreground/background border ($0 \leq d \leq 1$), a is the amplitude of the scaling, and σ specifies the width of the slope of scale values. The amplitude of the scaling is $a = 2(1 - C)$, for $C \leq 1$, otherwise no enhancement is required. The width of the slope, σ , is set according to the image size ($\sigma = 0.5$ for all our images as they have similar size). The border is obtained by partitioning the HDR

luminance Y into two segments using *K-means* image segmentation. A more elaborate image segmentation technique can be used for in challenging cases, for instance, when a simple segmentation returns unadjoining regions or when regions do not meet the image boundaries. Example m values are shown in Figure 4.4(b).

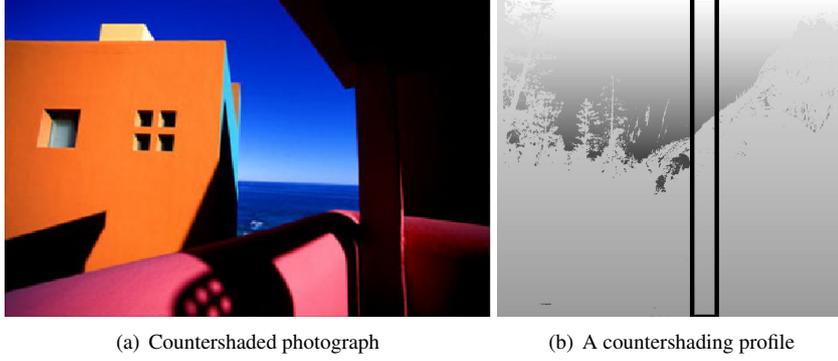


Figure 4.4: Pete Turner uses countershading to increase global contrast appearance (a). A global contrast signal (large-scale Cornsweet contours) along border of foreground and background (b).

Detail Visibility Restoration is a per-pixel operation that directly increases the visual contrast between detail pixels and their surrounding neighbourhood, thereby improving their salience in the image. This operation works much like modulating a base signal with a detail signal in image processing. To create contrast by chroma scaling, HDR high-frequency pixels are extracted with a low-pass bilateral filter and their chroma is increased in proportion to the detail visibility measure $\Delta\bar{T}^*$ defined in [Smith *et al.*, 2006]. Detail visibility scale values m_D are determined by relating the JND values of $\Delta\bar{T}^*$ to $\Delta C_{uv}^*(I', I) = |C_{uv}^*(I)(m - 1)|$. Through informal experiments, one JND is found to be approximately 6.89 ΔC_{uv}^* units.

$$m_D = \pm \left(\frac{6.89 \bar{T}^*}{C_{uv}^*(I)} + 1 \right) \quad (4.5)$$

In this technique, enhancement is not done by saturating one side and desaturating another, but instead by gradually increasing colourfulness on both the sides of the contour. A global chroma increase is commonly performed to improve the overall perceived quality of images [Fedorovskaya *et al.*, 1997]. Since humans favour increased saturation, the distortion approach uses positive scale values m_D . This does not in fact create a chromatic Cornsweet contour, because the chromatic peaks are not in opposition. However, in practice, the double-sided chromatic increase succeeded in making details more salient because contrast between vibrant colours appear higher than between less colourful colours.

Since tone mapping involves a tradeoff between detail preservation and global contrast compression, one restoration technique will be prominent. When both distortions are present, they will be slight, allowing the combination of both restorations. In these cases, it is best to begin with detail restoration so as not to disrupt the effect of global contrast restoration.

4.3.3 Unsharp Method

This newer approach fits into the unsharp masking paradigm, and has the advantage of being both automatic and single-step. Instead of using the term global contrast and detail visibility, the approach is aligned to humans' processing of spatial frequencies as described in Section 2.1.3, and restores *base layer* and *detail layer* contrasts with a unified contrast signal,

The contrast signal $C(C^*, Y, y)$ contains both the lost base and detail contrasts, and is derived by comparing the original and tone mapped luminances, denoted as Y and y respectively. Unsharp masking is applied to both the chromatic channels, and the gain values λ_{u^*} and λ_{v^*} are the amount of chromaticity measured by u^* or v^* , so $\lambda_{u^*} = u^*$ and $\lambda_{v^*} = v^*$. Using chroma channels as the gain values scale the contrast signal into the correct chromatic value range. Thus the enhanced tone mapped image $\mathcal{U}(I_{LUV})$ is defined as

$$\mathcal{U}(I_{LUV}) = [L^*, u^* + \lambda_{u^*} C(C^*, Y, y), v^* + \lambda_{v^*} C(C^*, Y, y)] \quad (4.6)$$

Since one image is in high dynamic range, contrast is measured by the logarithmic ratio of luminance between a single pixel and its averaged neighbourhood, as was used by [Mantiuk *et al.*, 2006a] to convert HDR luminance to contrast (Equation 2.1). Replacing the ratio by a difference of logarithms transforms the calculation into a Difference of Gaussians approach which estimates the Laplacian of the log luminance intensities, as described in Section 2.2.1. Since log luminance approximates brightness, this approach effectively measures local brightness contrast, which is better suited for comparing HDR and LDR images than absolute luminance contrast. Using a low-pass Gaussian filter with radius σ , contrast is determined as follows for HDR Y and LDR y :

$$C_{LogRatio}(Y) = \log_{10} Y - \log_{10} Y_{\sigma} \quad C_{LogRatio}(y) = \log_{10} y - \log_{10} y_{\sigma} \quad (4.7)$$

However, these definitions only measure fine high-frequency contrasts. To correlate to human viewing, the original images are split into their details and base layers. The base layer is constructed to contain the large scale brightness differences. Bilateral filtering is used in the construction since it preserves the position and magnitude of its boundaries. This is important, because these large scale boundaries will then determine the base contrasts to be enhanced, and should therefore not drift, as their enhancement must fall very near to their location in the LDR image. Two base luminance layers Y_B and y_B are created from Y and y by first applying a small Gaussian blur to remove noise, followed by a bilateral filter. This works well to preserve large transitions and their contours while removing smaller local transitions, as shown for the ‘‘Strasbourg’’ image in Figure 4.5.

The polarity of a contrast signal determined solely by $C_{LogRatio}(Y)$ and $C_{LogRatio}(y)$ represents the direction of brightness changes, and not the direction of chromatic change. Since it is destined to be added back to the chromatic channels, the polarity of the signal should reflect chromatic polarity. For instance, the darker side of the contour may be more chromatic, thus the darker side's chromaticity should be increased, not vice versa. Therefore, the polarity of the contrast signal will be controlled by the sign of chromatic contrast, which is defined by the HDR chroma C^* as:

$$C_{DoG}(C^*) = C^* - C_{\sigma}^* \quad (4.8)$$

The details layer contrast signal C_D is the difference between HDR brightness contrast $C_{LogRatio}(Y)$ and LDR brightness contrast $C_{LogRatio}(y)$. Although the values are



Figure 4.5: Luminance of the LDR image y , and the estimated base layer from HDR luminance Y_B (values are clamped to displayable range).

not perceptually meaningful, they reasonably approximate the log function for bright image regions, where the Weber law holds. Combining the magnitude of luminance contrast differences and polarity of chromatic contrast from the HDR image, the final details contrast signal is defined as:

$$C_D(C^*, Y, y) = \text{sign}(C_{DoG}(C^*)) |C_{LogRatio}(Y) - C_{LogRatio}(y)| \quad (4.9)$$

Using the base layers Y_B and y_B , and the chroma of the HDR base layer C_B^* , global contrast losses occurring between the predominant image sections can be estimated by C_B . In the same manner of defining detail layer contrast C_D , base layer contrast signal is defined as:

$$C_B(C^*, Y, y) = \text{sign}(C_{DoG}(C_B^*)) |C_{LogRatio}(Y_B) - C_{LogRatio}(y_B)| \quad (4.10)$$

Combining the detail and base contrast signals, the final contrast signal is defined as:

$$C(C^*, Y, y) = C_D(C^*, Y, y) + C_B(C^*, Y, y) \quad (4.11)$$

Due to processing the base layer, its average intensities become smaller than in the original image. However, because the base layer contrasts are also contained in the detail layer, no scaling is required, as the addition of the two layers' contrasts provides sufficient compensation for intensity lost due to bilateral filtering. The representation of the contrast signal in terms of base and detail contrast loss unifies the enhancement in an automatic way which is capable of driving both global and local contrast enhancements.

4.3.4 Resulting Images

The resulting enhanced images maintain a natural quality, and as such, some enhancements can be subtle. For this reason, and because of colour infidelity in print, the results are best visualized on screen. Results from the distortion metric based approach and the newer unsharp masking approach are shown alongside each other. There are visible differences between the two approaches' effects, yet the simple, automatic unsharp method can be seen to achieve the same quality of contrast restoration.

Included in the results are images from the following global (spatially uniform) *gamma correction* tone mapping operator ($\gamma = 2.2$), and the *photoreceptor* tone mapping operator, also global [Reinhard and Devlin, 2005]. The *gradient domain compression* operator [Mantiuk *et al.*, 2006a] and *bilateral filtering* [Durand and Dorsey, 2002]

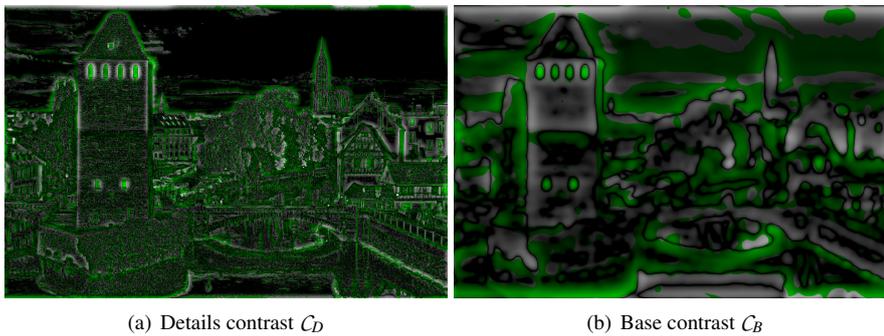


Figure 4.6: Visualization of the details and base contrast signals before they are combined. Green values represent negative values, white are positive values. The brightness of both represent the contrast to be added, so black values are regions where no enhancement is made.

were the local (detail preserving) algorithms that were used. The tone mapped LDR images were obtained either from the authors of these methods or by using publicly available implementations from [pfsTMO, 2008]. Tone mapping parameters were fine tuned whenever default values did not produce satisfactory images.

A tone mapping from *bilateral filtering* [Durand and Dorsey, 2002], shown in Figure 4.7(a) exhibits light global contrast loss of $C = 0.8747$. While the effect of countershading to restore global contrast (Figure 4.7(b)) and contrast restoration by the unsharp approach (Figure 4.7(d)), is subtle in print, on screen the increased chroma contrast at the horizon serves to emphasize the separation between the sky and the mountains. As can be seen in Figure 4.7(c), chroma is increased behind the rock, where the ground meets the horizon, and on the top of the rocky ledge. It is decreased (green values) on the rock face and at the clouds. This example shows that the base layer contrast calculated in the unsharp method is a very capable tool for measuring larger scale contrast losses.

The Strasbourg image, Figure 4.8(a), results from the gradient method tone mapping algorithm. The LDR version appears flattened, and the differences between sky, city and water are much less striking than in the HDR. The distortion method uses a global contrast loss ($C = 0.47626$) which results in a dramatic chromatic increase of the trees and topmost buildings, Figure 4.8(c). The unsharp masking version increases chroma at the horizon much like the manually segmented distortion method (Figure 4.8(b)). The contrast signal identifies the major elements of the image and creates a novel enhancement in the water's reflections, making less of a pronounced chromatic shift along the horizon, but nonetheless giving the scene a greater sense of depth and clearer separation between foreground and background, Figure 4.8(d).

Detail visibility greatly decreases in this poor quality LDR image resulting from simple gamma correction, Figure 4.9(a). This didactic result exemplifies how our enhancement technique reintroduces details and chromatic information into areas where they have been lost, in this case, enhancing the bleached sky surrounding the sun, Figures 4.9(c) and 4.9(d). The unsharp contrast signal shows how both details and base layers contribute to defining where the chromatic Cornsweet contours will be introduced, Figure 4.9(b).

The original LDR version of the café image tone mapped by *photoreceptor* [Reinhard and Devlin, 2005] is shown in Figure 4.10(a). In this example, the added chro-

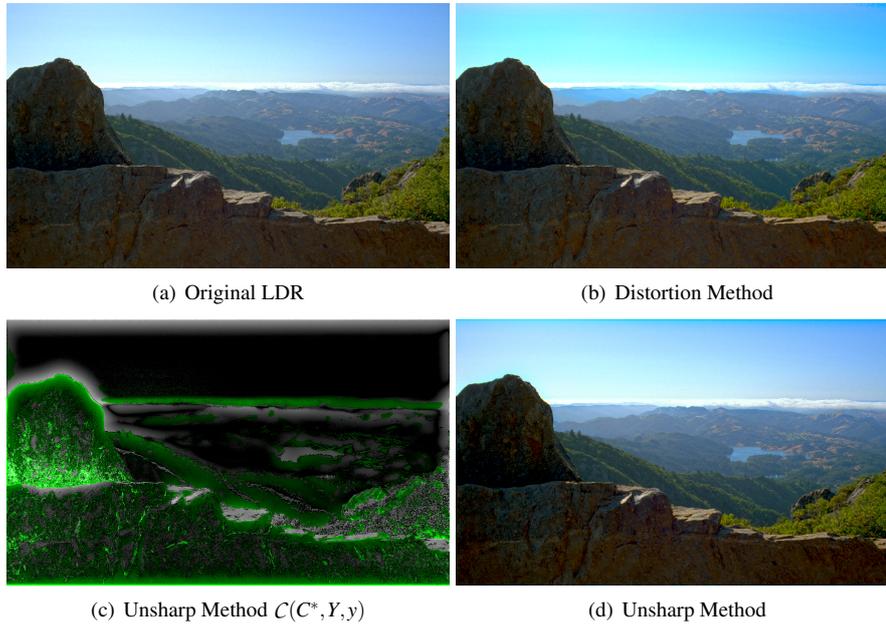


Figure 4.7: Global contrast enhancement of *bilateral filter* tone mapping (a) results in enhanced contrast (b&d) that is largely global in nature. The unsharp approach contrast signal shows where chromatic Cornsweet contours occur, and are then used to restore contrast to contrasts between large-scale regions (c).

matic contrast makes the details in the outdoor areas have closer appearance to the original HDR, Figures 4.10(b) & 4.10(c). The unsharp method has the advantage of adding chromatic contours in the foreground of the scene as well, due to the definition of the base layer. Figure 4.11(a) shows the contrast signal used for enhancement. Two enlargements shown in Figures 4.11(b) & 4.11(b) show the difference between the LDR image and the unsharp restored version, which has a much more detailed distant landscape and salient flower details as existed in the HDR original before tone mapping.

4.4 Summary and Discussion

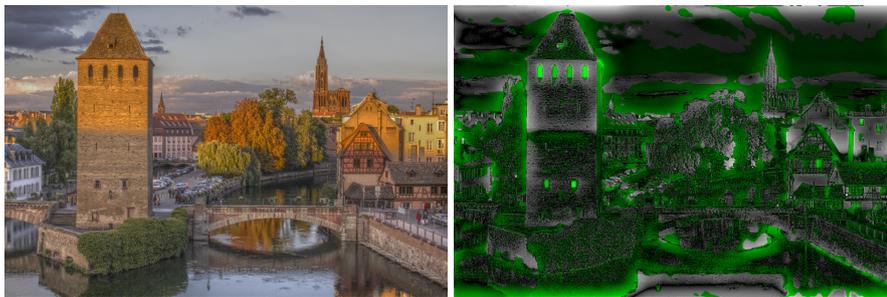
A related approach that was published after the distortion approach for Beyond Tone Mapping [2006] restores lost contrast by modifying the LDR luminance directly [2007]. The approach, named *Adaptive Countershading* uses the multi-scale luminance contrast of an HDR image to enhance the tone mapped version. The multi-scale contrast is determined using an uncompressed Laplacian pyramid, and contrast loss is measured as the ratio of LDR contrast and HDR contrast, as determined by the Weber contrast measure (Equation 2.2). So the contrast to restore is defined as $1 - \frac{C_{Weber}(y)}{C_{Weber}(Y)}$. It then incorporates models of visual masking of the Cornsweet illusion to prevent over-shooting the enhancement. Since visual masking is modeled, larger contrasts are allowed since the effect is weaker in textured regions. If the contrast restore measure is predicted to cause artifacts, it is reduced.

Consider the two approaches at a single spatial frequency scale, and disregard the over-shooting prevention. The remaining differences between the Adaptive Countershading approach and the unsharp masking Beyond Tone Mapping approach from Section 4.3.3 are in the contrast signal calculation and LDR channel to which it is added. Translated to the unsharp masking framework, Adaptive Countershading would be described as:

$$\log_{10}(\mathcal{U}(y)) = [\log_{10}y + (1 - \frac{C_{Weber}(y)}{C_{Weber}(Y)}) * (\log_{10}Y - \log_{10}Y_{\sigma})] \quad (4.12)$$

The most evident differences are that the enhancements are applied to luminance in one case, and chroma in the other. The next observation is that $\mathcal{U}(y)$ enhances luminance in log space by adding the HDR brightness contrast (measured by the Laplacian), then subtracting some amount of that HDR brightness contrast to prevent over-enhancement. The amount to subtract is controlled by the ratio of LDR luminance contrast over HDR luminance contrast. The common aspect is that both approaches determine the Cornsweet contours by calculating the Laplacian of HDR brightness, which is approximated by log luminance. To add these as chromatic Cornsweet contours, the brightness contrast is used directly to scale chroma channels u^* and v^* , since these are not in the same range as lightness, and do not have a logarithmic behaviour. Adding contrast with chroma is not as powerful an enhancement tool as adding luminance contrast directly, however, as shown in the results, it presents a good approach for luminance-preserving contrast enhancement.

The main contribution of the Beyond Tone Mapping approach presented here is that it provides a way to create iso-luminant Cornsweet contours, and extend the technique of unsharp masking directly to the colour channels. The approaches do indeed increase colourfulness of regions where contrast has been lost due to tonemapping, and show that with a careful and efficient adjustment of chroma, tone mapped images can better depict the scene captured by the original HDR image.



(a) Original LDR

(b) Unsharp Method $C(C^*, Y, y)$ 

(c) Distortion Method

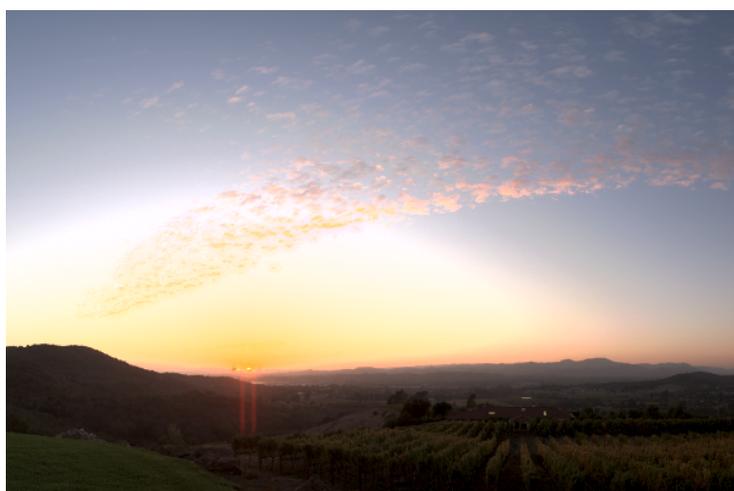


(d) Unsharp Method

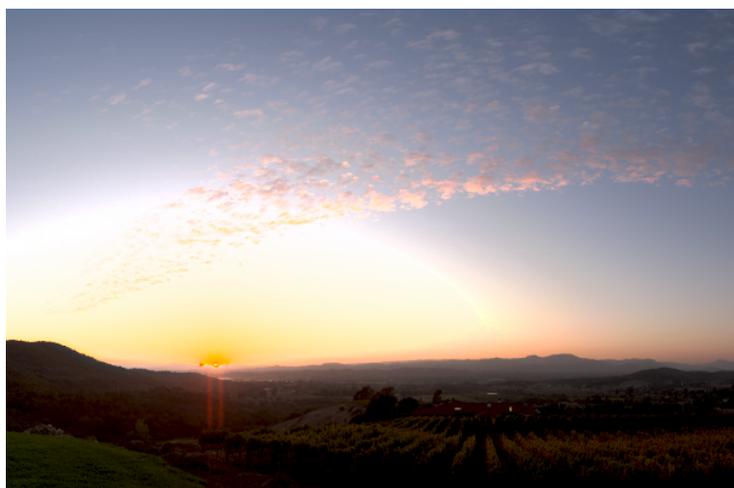
Figure 4.8: The Strasbourg image resulting from a gradient method tone mapping (a). Global contrast is enhanced by countershading using the distortion method (c). The unsharp masking approach is more subtle, but adds chromatic contrast in other areas of the image, such as the water's reflection (d). The result is that the higher contrast created with chroma helps to evoke a greater sense of scene depth and vibrancy.



(a) Original LDR

(b) Unsharp Method $C(C^*, Y, \gamma)$ 

(c) Distortion Method



(d) Unsharp Method

Figure 4.9: Gamma corrected LDR image $\gamma = 2.2$ (a), the unsharp method contrast signal shows where contrast is lost (b) and enhanced LDR images (c) & (d) with chroma added to restore the clouds and sun.



(a) Original LDR



(b) Distortion Method



(c) Unsharp Method

Figure 4.10: Detail restoration: Café image resulting from *photoreceptor* tone mapping (a), with chromatic contrast restoral by the distortion method (b) and the unsharp method (c). Notice that the flowers, chairs and umbrellas are more visible and the distant landscape contains more details and depth.



(a) Unsharp Method contrast signal $C(C^*, Y, y)$



(b) Trees close-up: original left, enhanced right



(c) Flowers close-up: original left, enhanced right

Figure 4.11: Contrast signal from the unsharp masking approach on the café image resulting from *photoreceptor* tone mapping (a) and enlargements of (b) & (c) showing enhanced depth and salient details after contrast restoration.

Chapter 5

Greyscale Conversion of Images and Video

The basic problem of greyscale transformation is to reproduce the intent of the colour original, its contrasts and salient features, while preserving the perceived magnitude and direction of its gradients. Photographs are now usually captured in full colour to allow greater post-processing abilities. However, printed material in newspapers, academic journals and textbooks contain black and white images due to printing costs. Some computer displays in medical or technical fields also display a limited range of colours, sometimes using only black and white for consistency. Many people prefer the look and style of black and white photographs, since it emphasizes the photo's content and appears more dramatic. Lastly, video and static images are frequently converted to a single grey channel for stylization and processing.

The problem of converting from a colour image to a black and white version requires the removal of colour while preserving the appearance of details and the overall brightness range. Artists frequently represent colourful imagery using only ink or pencil. As evidenced in Section 5.1, they make use of specially placed gradients or contours to overcome a limited palette.

Surveyed in Section 5.2, recent approaches solve discriminability constraints to determine the grey values, producing images in which the original colour contrasts are highly discriminable. For some, like those specializing in the conversion of business graphics, the goal is discriminability of all differently coloured regions. For others, the goal is to recreate the appearance of the original. The important characteristics when measuring the goodness of a greyscale conversion algorithm are the brightness ordering of difference colours, the new dynamic range and the local image features.

5.1 In Art: Contours Overcome An Achromatic Palette

One of the first challenges in drawing or painting after deciding how to map the 3D world onto a 1D surface, is deciding how to map a wide colour range to a more limited, sometimes achromatic, palette. In the example presented by Floyd Ratliff, a Chinese vase is decorated without colour [Ratliff, 1965]. The artist instead creates a Cornsweet contour to give the illusion of a bright sun in an equally bright sky, Figure 5.1(a). The colour differences between the sun and sky can not literally be represented, so the

contour creates the impression that the light source is distinguished from the sky while keeping both source and sky nearly isoluminant.

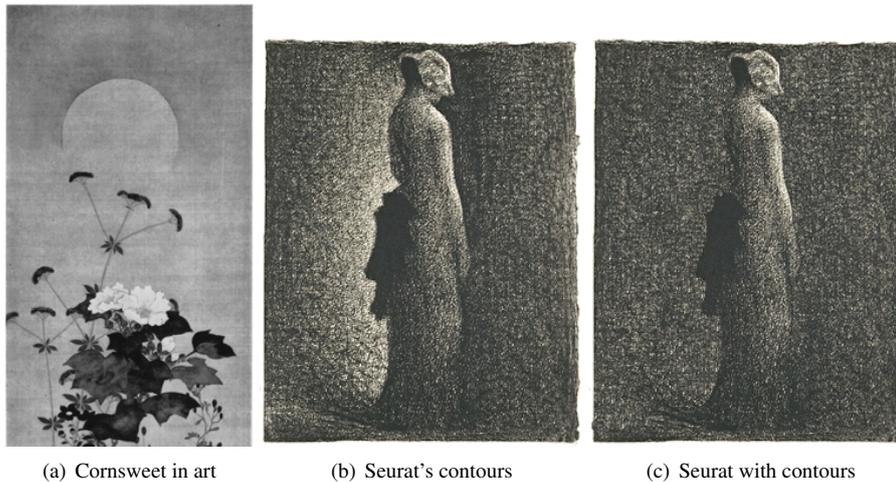


Figure 5.1: The decoration on a Chinese vase creates a bright sun with a Cornsweet contour (a). In *Le Noeud Noir*, Georges Seurat introduces contour effects to enhance overall contrast and emphasize the woman figure against the background (b) & (c).

A more global use of the Cornsweet border effect is seen in Georges Seurat's *Le Noeud Noir*. Here, the artist introduces contour effects to create a sense of greater contrast over the whole image and emphasizes the separation of the woman figure from the background, Figure 5.1(b). For the background, Seurat uses the Cornsweet contour effect to draw the contrast he sees, or wishes to be seen, since such a luminance gradient clearly would not arise from simple physical lighting. Consider the stark reduction in contrast when the contour effect is removed by replacing the gradient with a uniform background, Figure 5.1(c). The figure and background blend together; the woman's dress and the wall become visually connected.

5.2 Survey of Greyscale Conversion Methods

There are a variety of printing and display solutions for converting images from colour to greyscale. The transformation consists of two interdependent tasks: a mapping that assigns a grey value to each pixel or colour, and a discriminability constraint so that the achromatic differences match their corresponding original colour differences. The most straightforward conversions perform an image independent global mapping from colour to a grey values. One way to do so is by desaturating the image to remove colour. Another is to choose the values of a single colour channel and discard the other two colour channels, which mimics the effect of a traditional colour filter. This filtering approach places emphasis on certain differences between some colours, but completely loses others. Standard image processing software like Photoshop and Gimp use a simple mapping from RGB values to luminance Y similar to $Y = 0.3R + 0.59G + 0.11B$. Global mapping does not consider the specific image properties, and does nothing to ensure that the resulting achromatic differences preserve the colour differences in the original.

More complicated conversion methods have been presented in the research fields of computer graphics, image processing and vision science. Each method designs a way to define the chromatic contrasts it strives to preserve during reduction from three dimensional colour space to one dimensional luminance space. A first technique is to preserve the contrast between different colours in the image. This is dependent on the size of the colour palette, and does not scale for complex or natural images. If there are many colours, then an approach is to preserve chromatic contrast between pixels or regions of the image. There are certain benefits and downfalls to each approach, namely complexity, consistence over a series of like images and amount of user control required.

When the image is expected to have a limited range of colours and luminance levels, an image-dependent mapping can be performed. For instance, to convert business graphics to greyscale, Bala et al. [2004] propose a mapping method in which the distinct colours of the image are converted to greyscale values according to a simplified lightness predictor that incorporates the Helmholtz-Kohlrausch effect, as defined in Section 2.3.3. After the colours have been globally mapped to a single dimension of lightness values, their discriminability is maximized. This is done by measuring the differences between adjacent pairs of lightness values, then respacing the values over the whole range of available greyscale values according to their relative colour differences. The respacing is similar to a histogram stretching approach to contrast enhancement. The approach is uniquely for graphics with up to 10 colours, and is not applicable to complex images.

Their short paper studying chromatic contrast for greyscale conversion [Bala and Eschbach, 2004] is a spatial approach and introduces local contrasts in CIELAB by adding a high-pass signal to the lightness channel (see Section 2.3.2 for definitions of colour spaces). Their approach proceeds essentially like unsharp masking, but the contrast signal combines the magnitude of a^* and b^* high-pass signals and the polarity of L^* high-pass, as depicted in Figure 5.2. To prevent overshooting in already bright areas, the contrast signal is locally adjusted according to the lightness contrast. The authors of this approach are vision and colour scientists working on printing and reproduction problems for Xerox, thus they approached the greyscale conversion as less of a theoretical optimization or dimensionality reduction challenge, and more of a depiction problem. However, the algorithm is susceptible to problems in chroma and lightness misalignment, and only a single frequency level of contrasts are added to the image.

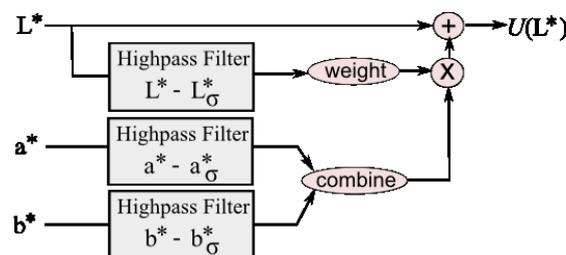


Figure 5.2: An unsharp masking approach to exaggerate chromatic contrast [Bala and Eschbach, 2004].

Gooch et al. [2005] created the popular *Color2Gray* algorithm, which finds grey values that best match the original colour differences through an objective function

minimization process. Original contrast between each pixel and its neighbours is measured by a signed distance, whose magnitude accounts for luminance and chroma difference and whose sign represents the hue shift with respect to a user-defined hue angle. Greyscale images resulting from this algorithm are presented in Figures 5.12, 5.13 and 5.21. It has $O(N^2)$ to $O(N^4)$ complexity, but a recent extension to a multiresolution framework by Mantiuk et al. [2006b] improves the algorithm's performance. These two approaches are dependent on the number of colours and image resolution. Rasche et al. [2005] propose a similar approach that finds the linear transform matching pairwise grey differences to corresponding colour differences. The best transform is found by minimizing an error function that can be evaluated over a smaller set of colours to alleviate computation costs. Images resulting from this algorithms are presented in Figures 5.15 and 5.23.

Grundland et al. [2007] construct a global continuous mapping that adds lost chromatic information to the luminance channel. Their algorithm achieves linear-time performance thanks to Gaussian pairing sampling which limits the amount of processed colour differences. In their custom (non-standardized) YPQ colour space, the colour differences are projected onto the two predominant chromatic contrast axes and are then added to the luminance image. This is similar to performing a principle component analysis (PCA) to reduce the dimensionality of the colour values. A saturation-controlled adjustment of the output dynamic range is adaptively performed to balance between the original range and the desired amount of enhancement. Their analysis of chromatic differences does not capture differences that are spatially distant, which can result in the same colour mapping to different grey values. This also means that original achromatic colours may be converted to substantially different grey values. Kuhn et al. [2008] present an approach that begins by quantizing the colours of the image then optimizes the quantization mass-spring system controlled by the chrominance of each colour. The final grey values for each pixel are then interpolated. The mass-spring approach enhances the contrast of the quantized grey values and preserves the value of achromatic colours in the original original.

Recently, Neumann et al. [2007] present a technique with linear complexity that requires no user intervention. It stresses perceptual loyalty by measuring the image's gradient field by colour differences in their *Coloroid* colour space. After discarding all gradient field inconsistencies, fast 2D integration determines the final grayscale image. Bloj et al. [2007] also formulate the greyscale problem as a gradient reconstruction in which the gradient field measures colour contrast. Reintegrating the gradient field determines a greyscale image whose luminance contrasts reproduce the original chromatic gradient field. A final tone mapping adjusts the resulting greyscale image so that its luminance histogram matches that of the colour image.

Calabria and Fairchild find that image lightness strongly affects perceived contrast, meaning techniques that can arbitrarily modify lightness, like approaches by Rasche and Grundland, may affect image appearance in an adverse way [2003b]. A greyscale ordering that contradicts the colours' luminance ordering also strongly impacts image appearance, yet in several approaches, ordering is subjective and arbitrary: the choice of hue angle in Gooch's Color2Gray can change all gradient directions, in Rasche's approach a user-defined threshold controls whether a colour is mapped to a darker or lighter value (see Figure 5.15), and in Grundland's approach ordering depends on the image and parameter choice of the colour sampling method.

The colour to grey methods discussed in Section 5.2 that depend strongly on local image content, colour palettes and user parameters are hindered in their perceptual accuracy. In the worst cases, images may suffer from detail loss, interpolation artifacts,

exaggerated dynamic range, arbitrary polarity over edges, extreme greyscale reordering and inconsistency among like images because each unique colour palette results in a unique greyscale mapping. For instance, image details and salient features may be lost by the choice of neighbourhood size in Gooch's Color2Gray or by unpredictable behavior in inconsistent regions of the gradient field in Neumann's approach (see Figure 5.21). Series of like images requiring comparison are printed in greyscale in catalogues and textbooks, and video is a sequence of like images. Converting image series and videos requires a greyscale transformation that is consistent and temporally coherent when applied in a frame-by-frame manner. The surveyed approaches that optimize based on the colour palette are not directly applicable to animation, since the colour palette is frequently modified and pixel correlations change quickly due to occlusion and disocclusion.

5.3 Apparent Greyscale

The following is novel work published as [Smith *et al.*, 2008] under the name *Apparent Greyscale: A Simple and Fast Conversion to Perceptually Accurate Images and Video*. The method treats greyscale conversion as a depiction problem whose solution is *perceptually accurate* version of the colour image that represents its psychophysical effect on a viewer. A perceptually accurate image is one that emulates both global and local impressions: it matches the original values' range and average luminance, its local contrasts are neither exaggerated nor understated, its grey values are ordered according to colour appearance and differences in spatial details are imperceptible. Strong perceptual similarity is particularly important for consistency over varying palettes and temporal coherence for animations, thus in addition to photographs, the approach is well suited to converting series of similar images printed textbooks and catalogues, the stylization of videos, and for display on monochromatic medical displays.

The Apparent Greyscale algorithm is a two-step greyscale transformation that combines a global mapping based on perceived lightness with a local chromatic contrast enhancement that adds Cornsweet contours. The success of the algorithm is empirical evidence that Cornsweet contours introduced through unsharp masking can be employed to overcome an achromatic palette. First, the grey values are mapped pixel-wise from each colour's apparent lightness, resulting in the reproduction of the original contrast and gradients. Second, the colour contrasts are measured in perceptual difference ΔE_{ab}^* and adjusted to maintain or improve discriminability with multi-scale unsharp masking. This twofold approach mimics aspects of the human visual system, which processes global attributes while simultaneously depending on local contrasts such as edges and surrounds.

5.3.1 Global Mapping to Apparent Lightness

It can be argued that the most accurate greyscale reproduction is that which represents accurate apparent lightness of the original colours and which preserves discriminability and ordering particularly in cases of isoluminant colours. Colour *lightness* is the best grey value, and provides the best ordering, because it is the HVS' achromatic response to a colour stimulus, measuring how bright a colour appears compared to an equally bright white. Section 2.3.1 provides an introduction to colour terms and the colour spaces that will be discussed below.

Colour studies show that lightness depends largely on luminance, but that colourfulness also contributes, as characterized by the Helmholtz-Kohlrausch effect (H-K effect), in which a colourful stimulus appears more light than a similar less colourful sample. The H-K effect has been identified as an important factor in greyscale mapping, and although it has been used for clipart greyscale mapping [Bala and Braun, 2004], no existing greyscale conversion for complex images explicitly takes it into account. The *global apparent lightness mapping* described here is independent of the original colour palette, incorporates the H-K effect so is sensitive to small differences even between isoluminant colours, and yields perceptually accurate gradient directions and an appropriate dynamic range.

As detailed in Section 2.3.3, the H-K phenomenon is predicted by a *chromatic lightness term* that corrects L^* based on the colour's chromatic component. Some models, like CIECAM02, account for many more complex colour appearance aspects, like surrounding colours, but are less suited to greyscale conversion due to their complexity and because most disregard the Helmholtz-Kohlrausch effect (the reader may refer to Table 17.1 in [Fairchild, 2005]). Of the standard recognized colour models, three suitable candidates for predicting chromatic lightness that incorporate the H-K effect: Fairchild's L^{**} [Fairchild and Pirrotta, 1991], and Nayatani's L_{NVAC}^* and L_{NVCC}^* [Nayatani, 1998].

$$L^{**} = L^* + (2.5 - 0.025L^*) \left(0.116 \left| \sin \left(\frac{H^* - 90}{2} \right) \right| + 0.085 \right) C^* \quad (5.1)$$

$$L_{NVAC}^* = L^* + [-0.1340 q(\theta) + 0.0872 K_{Br}] s_{uv} L^* \quad (5.2)$$

$$L_{NVCC}^* = L^* + [-0.8660 q(\theta) + 0.0872 K_{Br}] s_{uv} L^* \quad (5.3)$$

These chromatic lightness metrics solve a key challenge in greyscale conversion because they predict differences between isoluminant colours. Figure 5.3 plots the lightness measured by each metric on a nearly equiluminant colour ramp. It can be seen that more variation occurs when the H-K effect is being predicted, compared to luminance-based L^* which predicts nearly equal lightness for all colours. Note that other colour pairs will map to the same greyscale value, but that these are predicted to be more similar than the isoluminant colours.

But which predictor is best suited to greyscale conversion? The L^{**} or L_{NVAC}^* have the advantage of being modeled on VAC data, which was gathered by asking observers to find a grey that matches a colour, and is thus akin to the greyscale conversion task. Moreover, in testing L_{NVCC}^* , its stronger effect maps many bright colours to white, making it impossible to distinguish between very bright isoluminant colours. For that reason, and by heeding Nayatani's advice that L_{NVAC}^* , instead of L_{NVCC}^* , should be used for predicting differences between isoluminant colours, L_{NVCC}^* is excluded from consideration [Nayatani, 1998].

Because they are both fit to VAC data, the behaviours of L^{**} and L_{NVAC}^* are very similar. Their differences stem from the data on which they are based, and the flexibility of the models. L_{NVAC}^* is based on both Wyszecki 1964 and 1967 data, theoretical arguments about H-K effect, and the effect of adapting luminance. The L^{**} model is based only on Wyszecki 1967 data and has a simpler treatment of hue which is likely responsible for the following characteristic: L^{**} of blue hues is much higher than L^* . This reduces the range of L^{**} values and makes its ordering differ significantly from both

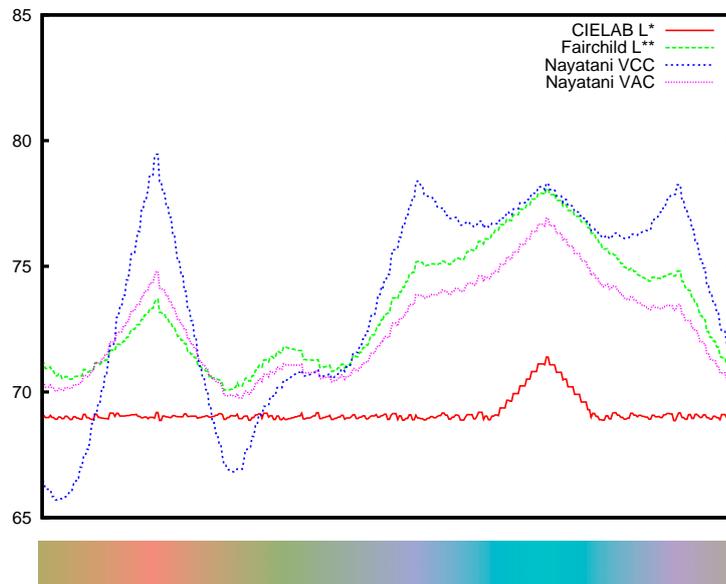


Figure 5.3: Lightness values from various H-K effect predictors applied to a spectrum of isoluminant colours, compared to CIE L^* .

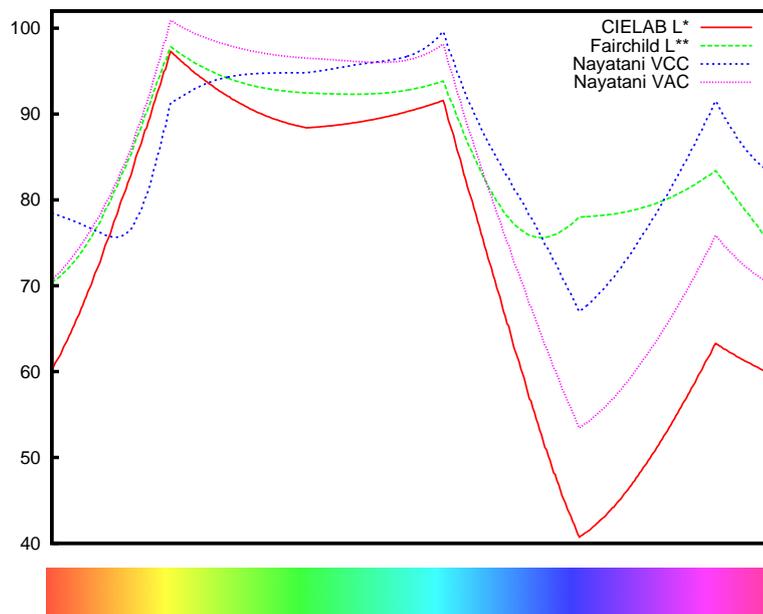


Figure 5.4: Lightness values from various H-K effect predictors applied over a full spectrum. L^{**} exhibits a small range and at blue hues differs from L^* .

$L_{N_{VAC}}^*$ and L^* , as shown in Figure 5.4. While the model fits the H-K effect perceptual data, its range reduction is problematic for greyscale conversion because colours with

different L^* become less discriminable, an observation shared by Bala [2004].¹ Therefore, L_{NVAC}^* is the most suitable H-K predictor to use as a global colour to greyscale mapping.

Using the Nayatani model $L_N^* = L_{NVAC}^*$, the global mapping to apparent lightness proceeds as follows:

$$I_{RGB} \rightarrow I_{LUV} \rightarrow I_{L_N^*} \rightarrow G \quad (5.4)$$

First, the colour image is converted to linear RGB by inverse gamma mapping, then transformed to *CIELUV* colour space because that is the space in which L_{NVAC}^* is defined. The apparent chromatic object lightness channel L_N^* is calculated according to Equation 5.2. Then L_N^* is mapped to greyscale Y values using reference white chromatic values for u^* and v^* . Finally, gamma mapping is applied to move from linear Y space back to a gamma-corrected greyscale image G . As shown in Figure 5.5 for several colour ramps, the mapping is continuous, there is no colour reordering, very little perceivable discrimination loss and the dynamic range is preserved.

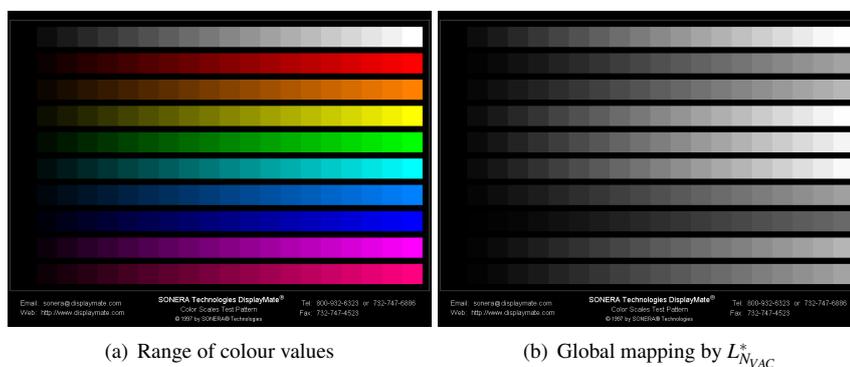


Figure 5.5: On a colour test image (a), a global mapping to apparent lightness using L_{NVAC}^* to greyscale G (b) preserves overall appearance and lightness ordering.

Due to the compression of a 3D gamut to 1D, L_N^* may map two different colours to a similar lightness, which then are quantized to the same grey value. This occurs only when colours differ uniquely by hue, which is very uncommon in natural images and well-designed graphics. Even for a very challenging image that comprises equiluminant colours sampled from [Neumann *et al.*, 2007], global mapping discriminates appropriately, predicting the H-K effect that makes a more colourful blue appear lighter than the duller yellow, as shown in Figure 5.6 (view original colours on a calibrated screen). Recall that the goal is perceptual accuracy: the resulting low contrast properly represents the low contrast of the colour image, and each unique colour is mapped to a unique greyvalue. By incorporating the H-K effect, the global mapping partially solves the problem of grey value assignment and appropriately orders colours that normal luminance mapping can not discriminate.

5.3.2 Local Chromatic Contrast Enhancement

The mapping from a 3D to 1D colour space reduces the overall difference between colours, jeopardizing discriminability. Humans are not too sensitive to this loss when

¹Bala uses $L_1^{**} = L^* + 0.143C^*$.

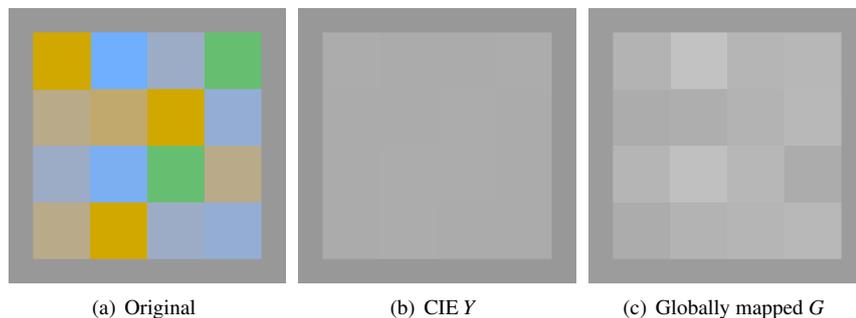


Figure 5.6: The global mapping based on chromatic lightness predicted by $L_{N_{VAC}}^*$ maps isoluminant colours to unique, properly ordered greyscale.

it occurs between spatially distant colours, but with adjacent colours it is immediately apparent, especially if an original contrast becomes imperceptible. To solve this problem, local contrast is enhanced until its magnitude emulates that in the original. The enhancement restores chromatic differences without overemphasizing luminance differences by adaptively increasing weak contrasts. Furthermore, the process is restricted so that the polarity over edges, overall lightness and colour ordering are preserved, thus maintaining perceptual accuracy.

Contrast adjustments are performed using the Laplacian pyramid that decomposes an image into n band-pass images h_i and a single low-pass image l [Burt and Adelson, 1983]. Recall that the Laplacian pyramid was introduced in Section 2.2.2 as a multi-scale way to measure local contrast, and the Laplacian is often the contrast signal which is added to an image during its unsharp masking Section 2.4.4. Laplacian pyramids are built for both I and G in *CIELAB* using a binomial coefficient filter of order 4. *CIELAB* is used because it is the recommended space for colour difference calculations. The h_i of each channel measures its local contrast, but as G contains no chromatic information, its local contrasts are contained entirely in its L^* channel.

The image reconstructed from the n band-pass images $h_i(G_{L^*})$ is the contrast signal used to enhance G_{L^*} . However, it must first be scaled to represent the amount of contrast to be added. At each scale in the Laplacian pyramid contrast signal, $h_i(G_{L^*})$ is adaptively increased by a perceptually-based amount λ_i , which measures the amount of contrast needed to match colour contrast $h_i(I)$. The enhanced greyscale image $\mathcal{U}(G)$ is computed by unsharp masking G_{L^*} as follows:

$$\mathcal{U}(G)_{L^*} = G_{L^*} + \sum_{i=0}^{n-1} k_i \lambda_i h_i(G_{L^*}) \quad (5.5)$$

where parameters k_1, \dots, k_{n-1} , $k_i \leq 1$ exist so that the spatial effect can be controlled according to amount of discriminability desired and the intended viewing conditions (image size and viewing distance).

The goal of gain factor λ_i is to measure the remaining chromatic contrast to be restored during the enhancement. Thus the chromatic contrast of the original image is measured using the ΔE_{ab}^* colour difference. Recall that ΔE_{ab}^* is defined in Equation 2.17 as the distance between any two colours in *CIELAB*. As a perceptually uniform difference, it allows comparisons between colour and grey differences in units of perceptual lightness. Thus, $\Delta E_{ab}^*(h_i(I))$ is the colour contrast between a pixel and its

neighbourhood, measured by:

$$\Delta E_{ab}^*(h_i(I)) = \sqrt{h_i(I_{L^*})^2 + h_i(I_{a^*})^2 + h_i(I_{b^*})^2}$$

Since the chromatic channels of G contain no contrast information, the chromatic contrast of G is roughly the contrast of its L^* lightness channel:

$$|h_i(G_{L^*})| \cong \Delta E_{ab}^*(h_i(G))$$

Using these two contrast measures, gain factor λ_i is defined as:

$$\lambda_i = \left(\frac{\Delta E_{ab}^*(h_i(I))}{|h_i(G_{L^*})|} \right)^p \quad (5.6)$$

The parameter $0 \leq p \leq 1$ is used to remap the λ values to a non-linear scale so that weaker contrasts, like those from isoluminant colours, can be enhanced without over-emphasizing stronger contrasts. To prevent the ad hoc nature seen in other approaches, the parameters provide flexibility without allowing uncontrolled changes to the image. Most importantly, by definition of λ , edge polarity can not flip, meaning the lightness order of adjacent regions can not be changed.

The effect of the Apparent Greyscale's local chromatic contrast adjustment is illustrated in Figure 5.7. Contrast is nearly below threshold between isoluminant regions in G , especially among the bottom row of colours, Figure 5.7(d). A basic sharpening of all contrast in G , Figure 5.7(e), does little to discriminate along that bottom row because there is no local contrast detected by the Laplacian calculation, Figure 5.7(b). However, with chromatic adjustment it is possible to lift these contrasts above threshold without over-emphasizing existing contrasts, as shown in Figure 5.7(f) so the resulting image better represents the original contrast. Notice that the Apparent Greyscale contrast signal contains far more information than the basic unsharp contrast signal, Figure 5.7(c).

As described in Section 3.4, the Laplacian of local contrasts can be used to create Cornsweet contours. So, it is helpful to think that the Cornsweet contours are defined by the contrast signal $\sum_{i=0}^{n-1} k_i \lambda_i h_i(G_{L^*})$. The parameter k_i scales the magnitude of each band, adjusting the height of the intensity ramps, and since it is performed separately for the different frequency bands, it controls the widths of the ramps in the resulting Cornsweet contours. Viewed in the context of unsharp masking, Apparent Greyscale proceeds as shown in Figure 5.8. The perceptual effect of the Cornsweet contours can be seen clearly in Figure 5.7(f), where nearly isoluminant regions are made to have different apparent brightness solely by the addition of the contrast signal.

The parameters k and p exist to provide flexibility, allowing users to tweak according to their preference for desired discriminability. This necessity is confirmed by [Connah *et al.*, 2007], who in comparing greyscale conversion methods, found that preference depended heavily on the content of the original image. The overall lightness is not altered because the number of subbands that may be enhanced is limited, preventing changes to the base layer low-frequencies (in practice $n \leq 4$ levels).

As previously mentioned, parameter p controls the remapping of λ so that weak chromatic contrasts may be enhanced without over-enhancing stronger chromatic contrasts. This is especially necessary for images with a wide distribution of chromatic contrast, namely those that contain salient isoluminant contrasts, of which Claude Monet and the pointillists were so fond. For this reason, the effect of adjusting parameter p is demonstrated on Claude Monet's *Impression: soleil levant, 1873*. Notice

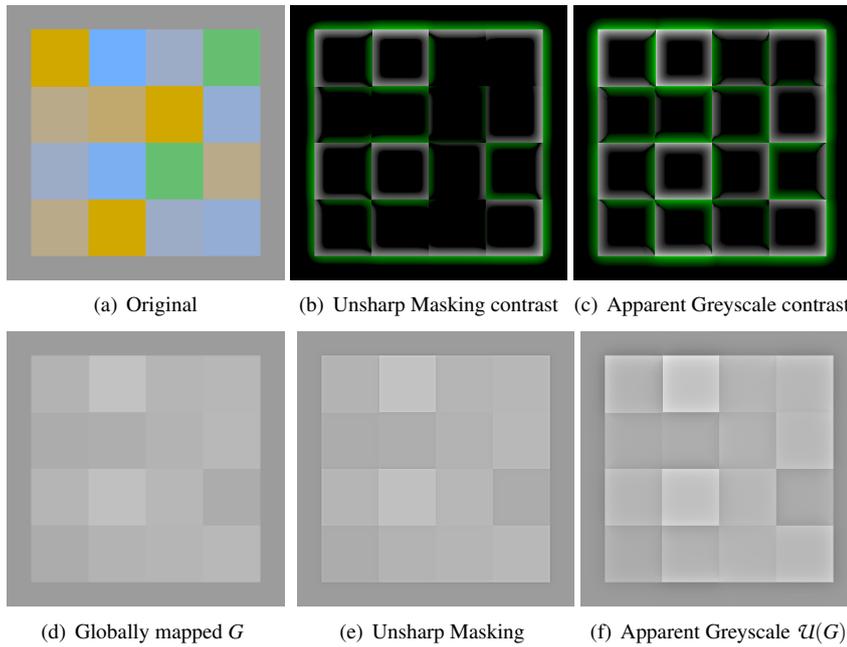


Figure 5.7: In the gain images, green values represent negative gain; $p = 0.3$ $k = \{1, 1, 1, 0.6\}$. Compared to the basic unsharp mask (e), Apparent Greyscale's chromatic enhancement (f) gains contrast where it is low in G and high in I .

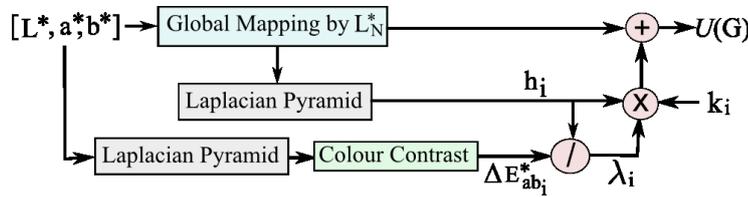


Figure 5.8: Organization of the Apparent Greyscale algorithm depicted in the unsharp masking framework.

that the boat and other luminance contrast are not enhanced. As desired, only chromatic contrasts are made more distinct, mainly in the sky, the sun and the watery reflections.

Parameters k_i control the width of contours that will be added and the overall strength of the enhancement. These parameter can be adjusted depending on whether the image is to be depicted at low or high resolution levels. At higher resolution, narrow contours will be visible, but at low resolution, the contours may need to be wider and stronger to have an effect on apparent contrast. Figure 5.10 shows the effect of reducing the impact of lower-frequency bands which results in reduced spatial extent of the enhancement. It also shows how a reduction of the highest-frequencies reduces the apparent contrast, and how a uniform weakening of all values the enhancement reduces discriminability.

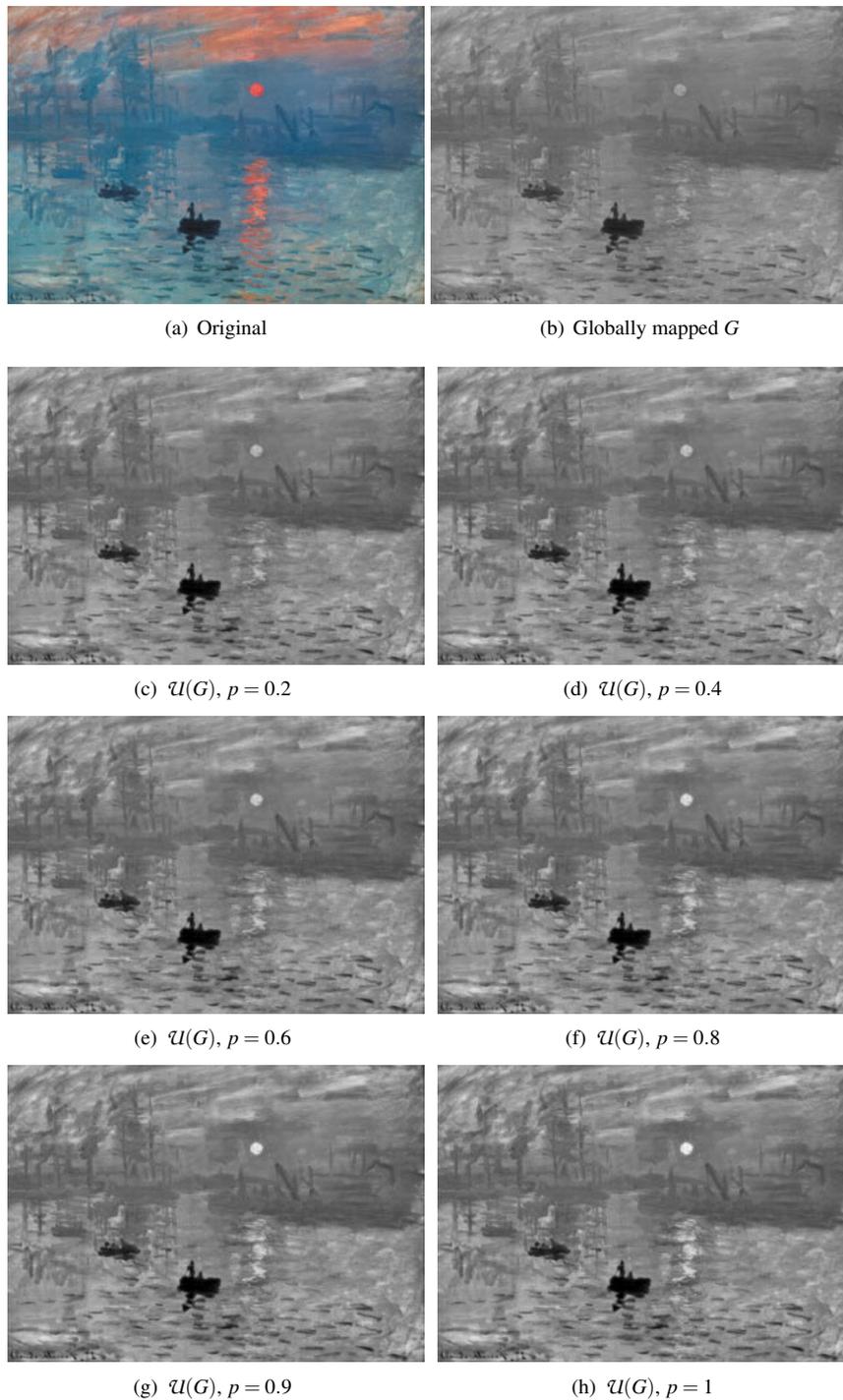


Figure 5.9: The effect of parameter p as shown on Claude Monet's *Impression: soleil levant, 1873*. Notice that the boat and other luminance discontinuities are not sharpened. Only chromatic contrasts appear, mainly in the sky and water's reflections. Parameter k is constant at $k = \{0.2, 0.5, 0.6, 0.4\}$.

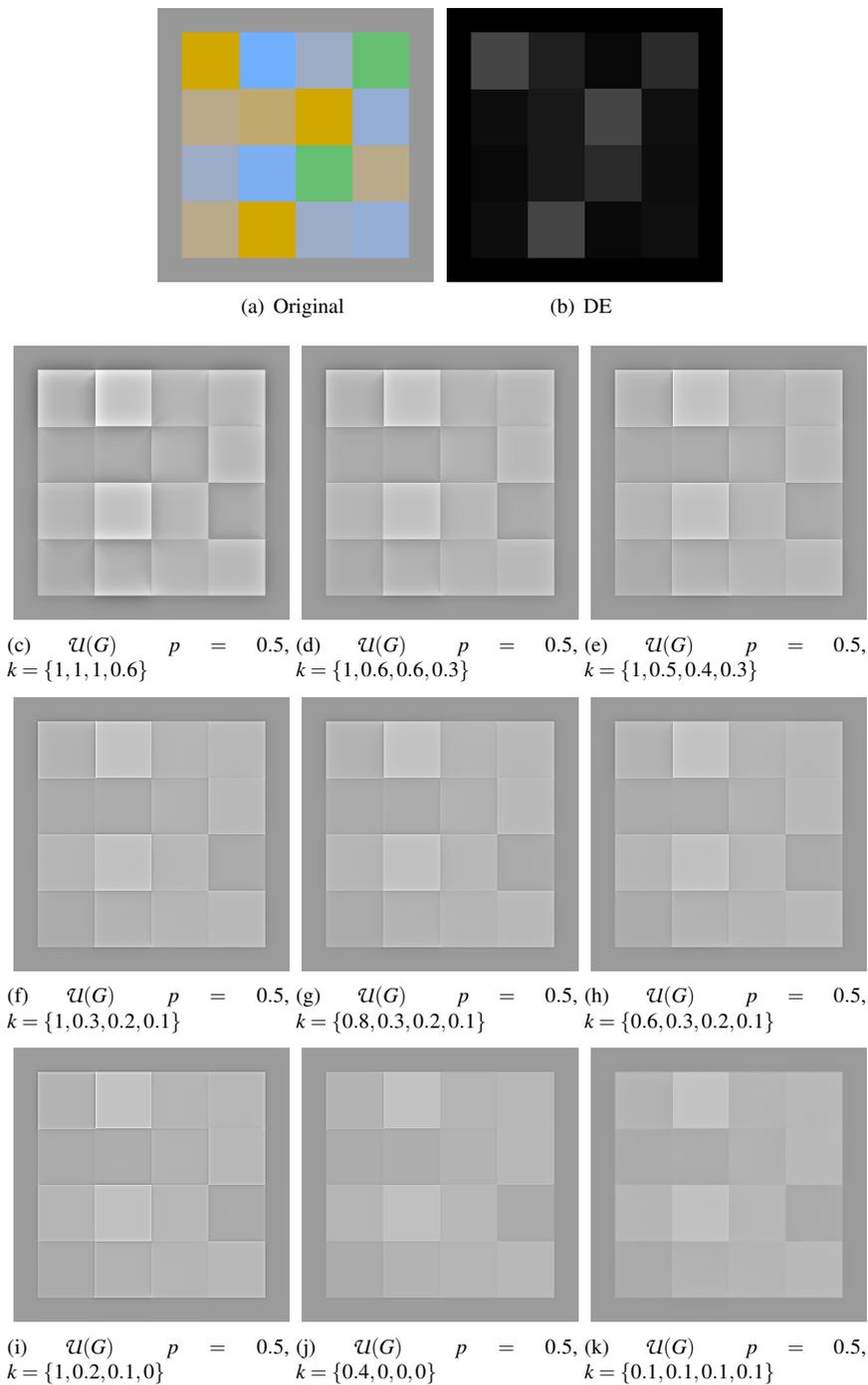


Figure 5.10: The effect of parameter k on an isoluminant pattern. The top row of greyscale images shows the effect of reducing the impact of lower-frequency bands, resulting in reduced spatial extent of the enhancement. The middle row shows a reduction of the highest-frequencies and the bottom row shows an overall weakening of the enhancement which reduces discriminability.

5.4 Applications of Apparent Greyscale

The results presented here strive for perceptual accuracy, and do not attempt to increase or exaggerate discriminability. Therefore, the effects are apparent, but subtle. For comparison, either the CIE Y channel or Gimp greyscale with a basic unsharp enhancement is presented, so that the reader is able to compare between images with matching overall sharpness. Additionally, the images presented here are for viewing on a calibrated colour screen (sRGB); for print, the resulting greyscale images should be mapped to the appropriate printer gamut. The following images make a convincing visual argument that the simple Apparent Greyscale algorithm yields comparable images to more complex approaches, and its linear runtime makes it suited to video processing and accelerated graphics hardware implementations.

5.4.1 Isoluminant Images

The approach works well in pathological cases, like the test blocks from Rasche et al. [2005] shown in Figure 5.11 and images from Gooch et al. [2005] shown in Figure 5.12. In these images, nearly isoluminant regions are distinguishable (parameters p and k , and thus n , are given with each image). There are cases where colours map to the similar same chromatic lightness, however, such collisions are infrequent and when they occur, it is because the colour model predicts that humans also have difficulty seeing a difference between the colours. Using the map image from Gooch et al. [2005], a strong enhancement is chosen so that the added Cornsweet contours are visible in the lake region, Figure 5.13.

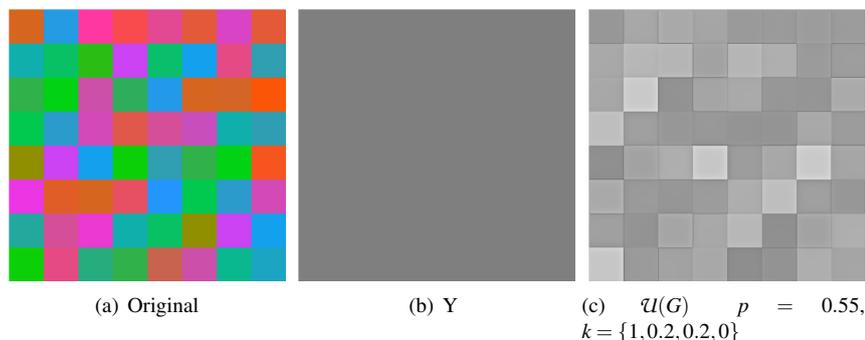


Figure 5.11: Rasche's isoluminant blocks are discriminated by Apparent Greyscale, and their ordering is consistent with the original colour ordering.

5.4.2 Image Sets and Video

To convert sets of similar images or frames of video, the resulting greyscale images must be consistent with each other. Figure 5.14 shows a set of flowers ordered by brightness. The greyscale images resulting from the Apparent Greyscale approach have a correct brightness ordering, the greyscale values of leaves and backgrounds are identical, and the lightness range is not exaggerated. The consistency problem of local adaptive approaches is illustrated on in Figure 5.15 by applying Rasche's approach with default parameters. This algorithm exaggerates the dynamic range of the images and sets each flower to bright white, thus losing the brightness order. The background greyscale values

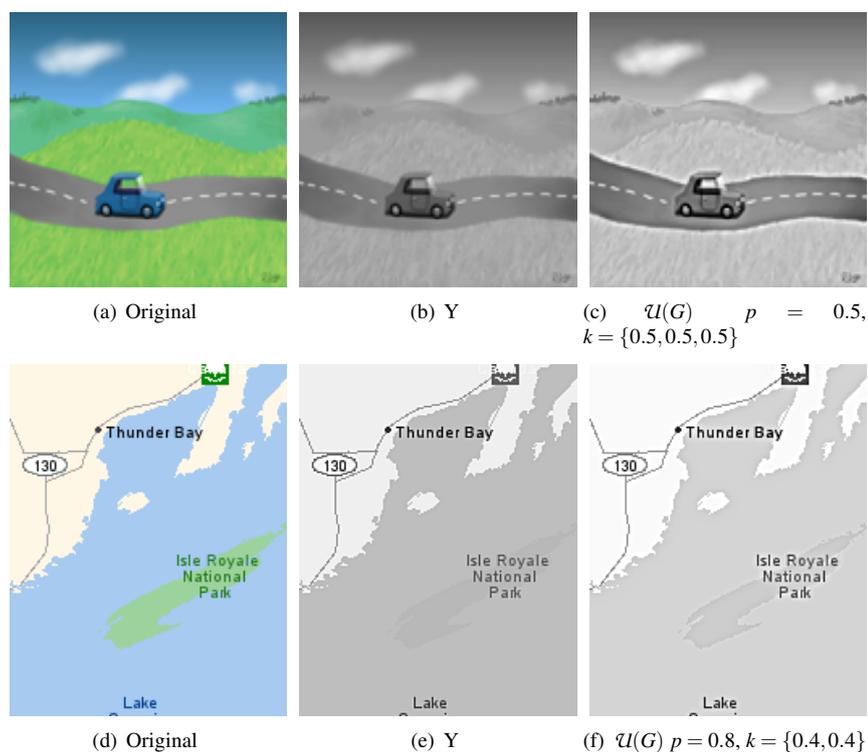


Figure 5.12: The greyscale conversion discriminates between colours even with very similar luminance.



Figure 5.13: In Gooch's map, the enhancement effect is intentionally made visible in the water because it has a pleasing look. The original colours are nearly fluorescent looking, but this may not appear in print.

are also dissimilar between the images, and the overall dynamic range has been greatly scaled.

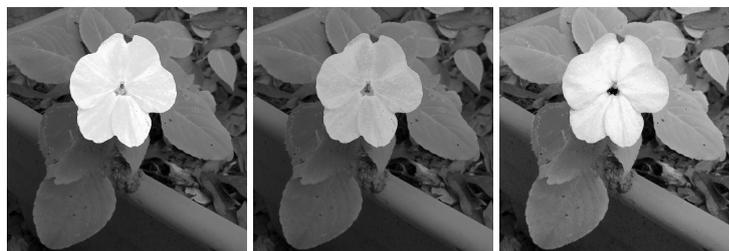
Because consistency is maintained, the conversion can be applied to video, as shown for a single frame in Figure 5.16 with constant parameters $p = 0.8$, $k = \{0.2, 0.8, 0, 0\}$. The local enhancement step risks introducing temporal inconsistencies, which is prevented by constant local parameters. The flowers become more visible without changing the overall video appearance and maintaining temporal coherence. The Apparent Greyscale converted videos are available at [Website, 2008b].



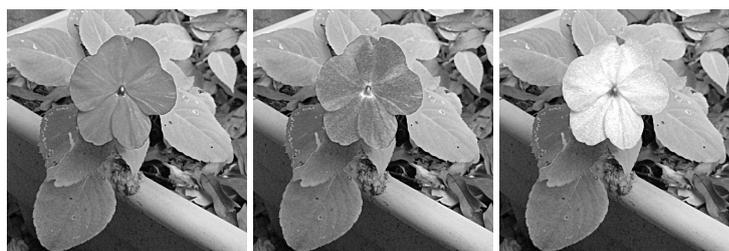
Figure 5.14: Consistent colours: A ‘bouquet’ of Impatiens, which the Apparent Greyscale approach orders without conflicts or rearrangements. Parameters are constant at $p = 0.5$ $k = \{0.5, 0.5, 0\}$.



(a) Original colour versions



(b) Converted with Rasche et al., default parameters



(c) $\mathcal{U}(G)$, $p = 0.5$ $k = \{0.5, 0.5, 0\}$.

Figure 5.15: Rasche (top), and Apparent Greyscale colour ordered conversion (middle), original colour (bottom).

5.4.3 Preserving Spatial Details and Salient Regions

Apparent Greyscale works well for converting highly complex images, producing depictions that are more perceptually accurate than similarly sharpened Gimp greyscale.

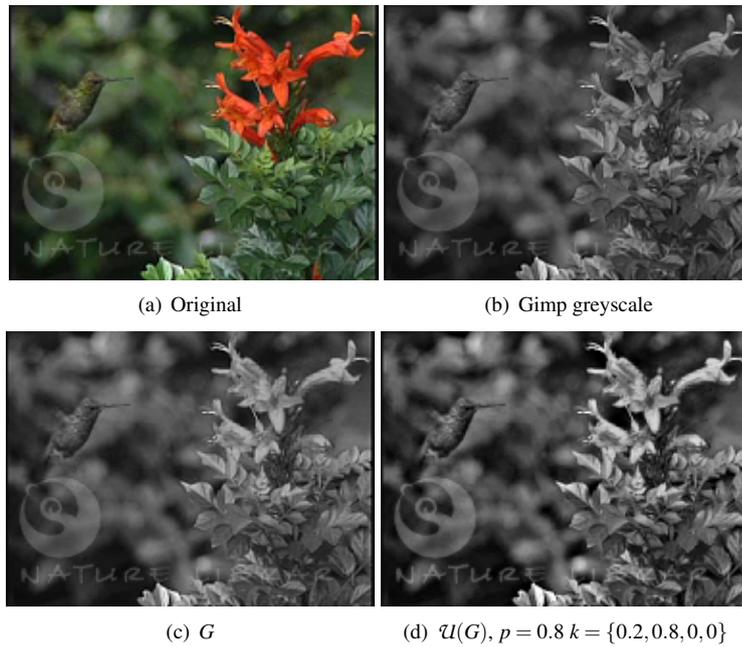


Figure 5.16: A frame from the hummingbird video. Source: www.naturelibrary.com. Visit [Website, 2008b] for the full video.

In Figure 5.17, the red fish and stone advance and the two orange fish reappear. In Figure 5.18, the hats are more bright and the furthest two are distinguished more easily. Apparent Greyscale is an excellent technique for photographs of real-world natural scenes as can be seen in Figure 5.4.3, where the red tulip becomes more salient, like it is in the original colour version.

Figure 5.20 presents a Monet painting showing that the Apparent Greyscale approach accurately determines that the poppies appear brighter than the field and the clouds brighter than the surrounding sky. Notice that besides the reappearance of poppies, the woman's figure and umbrella become visible and the little girl is once again holding a flower.

Local optimization approaches to greyscale conversion may change the spatial content of the image. For instance, a photograph of Claude Monet's *Impression: soleil levant*, 1873 captures paint strokes, blurry atmospheric effects in the painting and his characteristic isoluminant colours, Figure 5.21(a). A direct conversion to luminance preserves all the spatial details and contrasts, but the isoluminant sun and sky blend together, Figure 5.21(b). The result from Gooch's Color2Gray, shown in Figure 5.21(c), dilates the sun and reflection and has a strong blurring effect. The resulting image from Neumann, Figure 5.21(d), masks details of the background structures, and alters the water's brightness, giving the impression of another light source. The global mapping of Apparent Greyscale properly predicts the chromatic lightness of the sun and its reflection, Figure 5.21(e). Together with the local enhancement step, the resulting image from Apparent Greyscale preserves the lightness of regions, the brightness of the sun, keeps all paint strokes visible, and when visually compared to original contains very few spatial modifications, Figure 5.21(f).

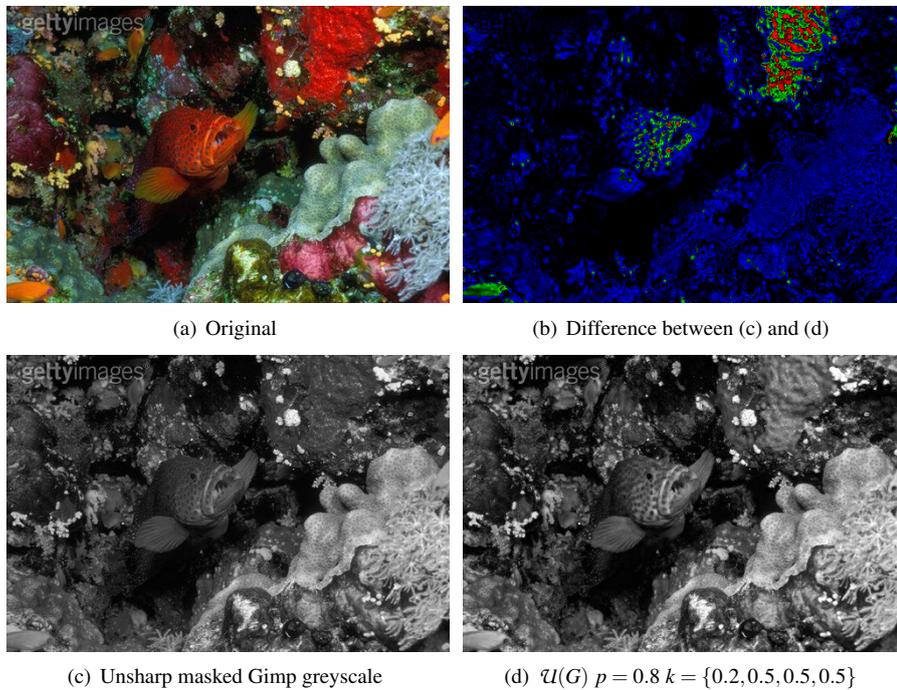


Figure 5.17: The greyscale version accentuates the red fish and stone, and restores salience to the two orange fish. Source: Getty Images.

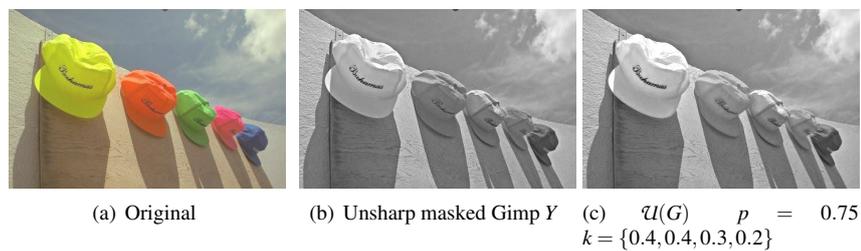


Figure 5.18: The extreme brightness of the hats is more apparent in $\mathcal{U}(G)$ than Gimp's greyscale which highlights the differences between the furthest two hats. Source: www.vischeck.com.

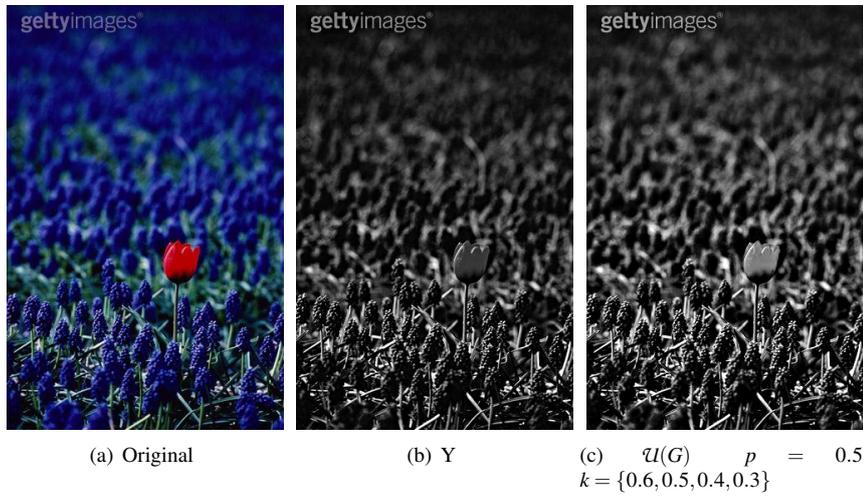


Figure 5.19: The red flower stands out from the background.

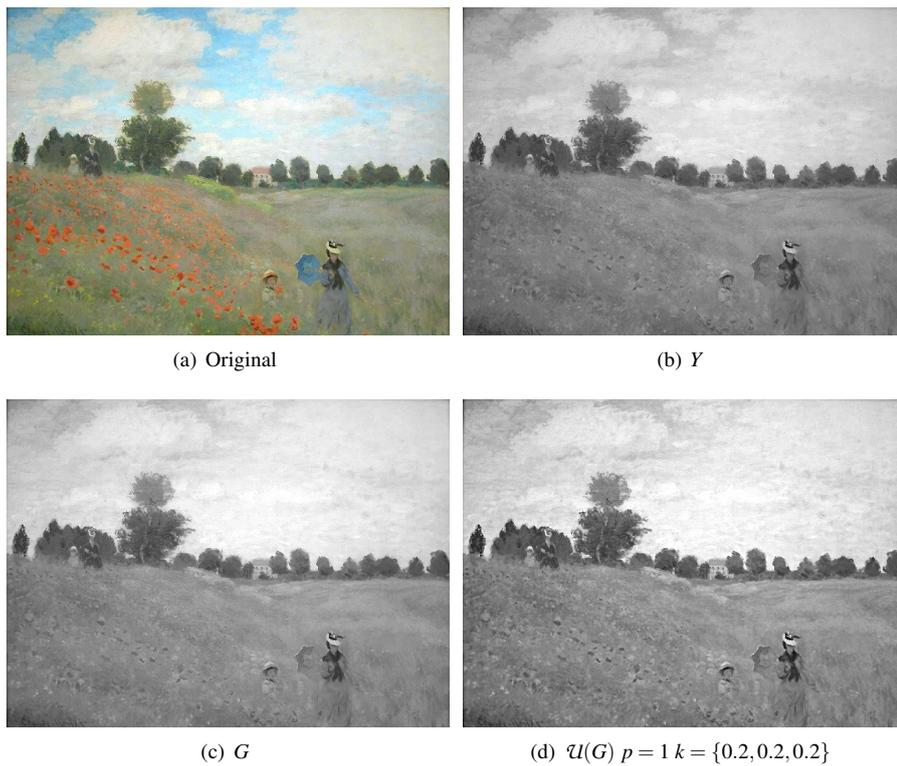


Figure 5.20: In Claude Monet's *Les coquelicots*, the greyscale converted image makes poppies that pop out like they do in the original, and separates the woman's blue dress from the green field, without enhancing the border between the trees and skyline.

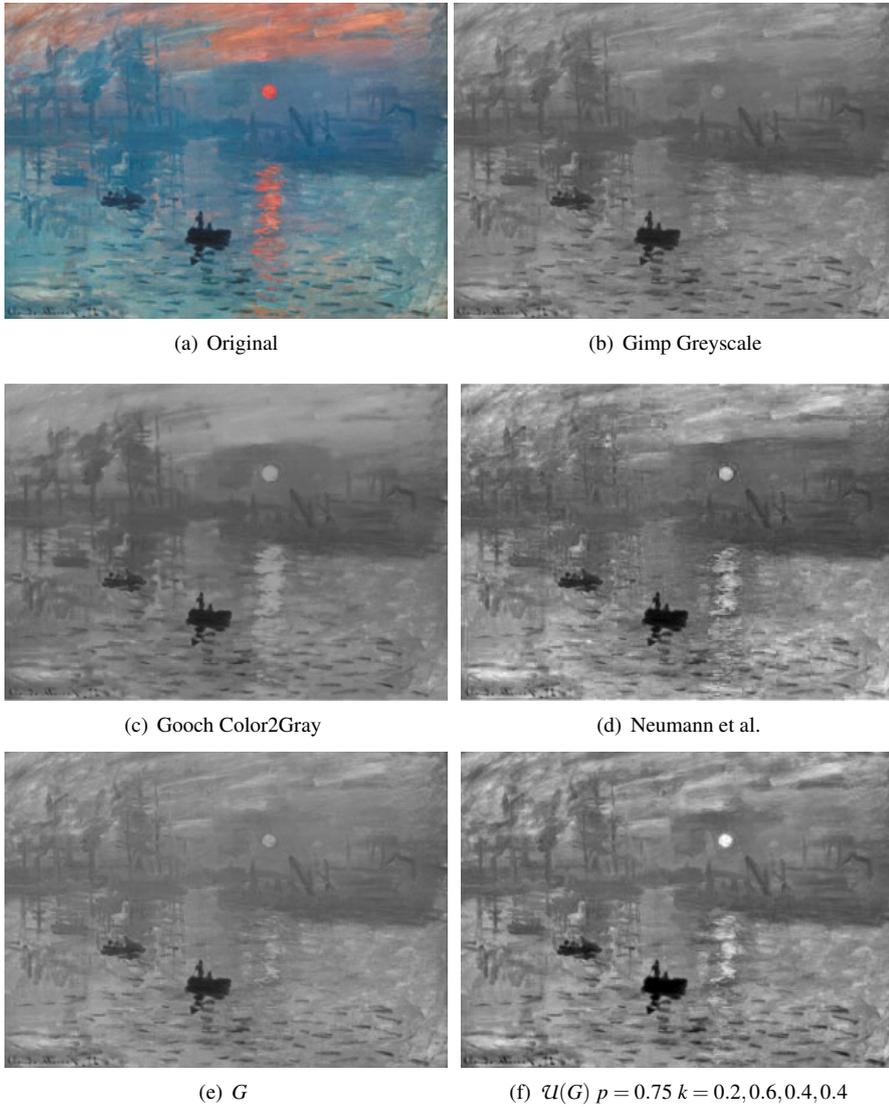


Figure 5.21: The impression from the Apparent Greyscale conversion (f) is more like the original because it preserves the paint strokes especially in the sky and background. Gooch's image (c) is strongly blurred with a dilated sun and Neumann masks the background and lightens the water (d).

5.5 Implementation and Limitations

The Apparent Greyscale algorithm has been implemented in Octave using the Colour Engineering Toolbox and a Matlab toolbox for Laplacian pyramids. Its runtime depends on image resolution and the speed of colour mapping and pyramid construction. The colour mapping could be significantly sped up with parallel computations or look-up tables. Times for both 1 and 4 storey pyramids computed on an Intel4 3 GHz CPU are as follows: The Impression Sunrise image (311×223 pixels) takes 1.8 or 3.2 seconds; the Impatiens image (570×593) takes 6.7 or 10.8 seconds; and the Hummingbird video (192×144) single scale conversion takes 136.3 seconds with 0.96 seconds for each of the 142 frames, no multi scale conversion was performed. The video is available for viewing at [Website, 2008b]. A single-scale GIMP plug-in, whose screenshot is shown in Figure 5.22, is provided at the website as well.

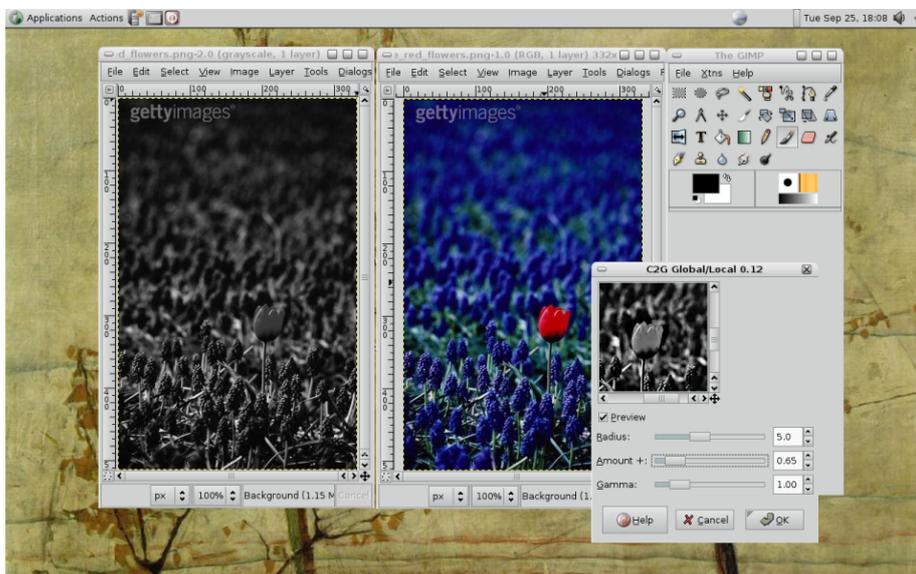


Figure 5.22: A single-scale version of the Apparent Greyscale approach has been implemented as a Gimp plug-in.

The main limitation of the Apparent Greyscale approach is the locality of the second step, since it can not enhance non-adjacent regions. In some cases, this may prevent the local approach to contrast restoration from attaining the same discriminability level as other optimization approaches. On a specialty image for testing daltonism shown in Figure 5.23(a), Rasche reach their goal of an exaggerated discriminability, Figure 5.23(b). Although Apparent Greyscale results in a perceptually accurate colour order and difference, Figure 5.23(e), the local chromatic contrast enhancement fails because the colours to be discriminated are separated by white boundaries, Figure 5.23(f). In such special cases where there is high desired discriminability, the chromatic lightness global mapping can be used and followed by a contrast stretch or automatic histogram equalization, thus keeping the benefit of colour ordering, Figure 5.23(g). Notice that compared to a Gimp greyscale with automatic histogram equalization, Figure 5.23(d), mapping according to chromatic lightness instead of luminance and then stretching contrast is more discriminable and its greyscale order is more perceptually accurate

because of its basis in colour studies.

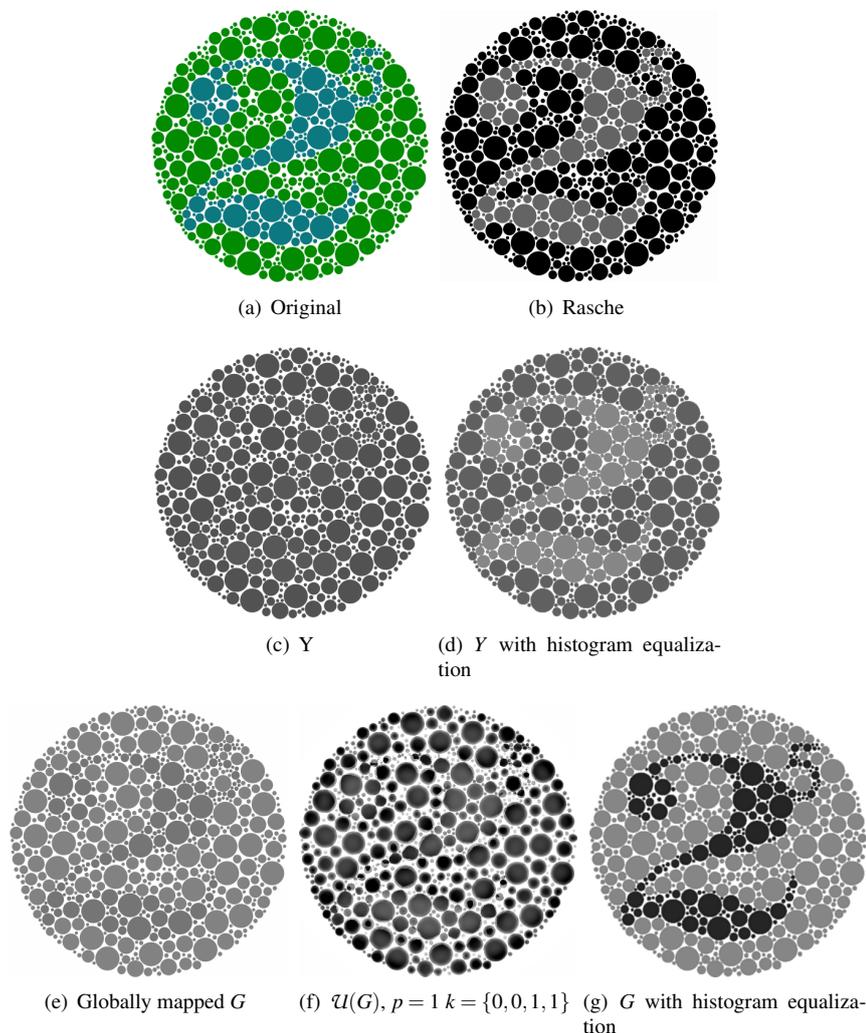


Figure 5.23: An example where local enhancement can not restore the lost chromatic contrast due to the separating white boundaries between the coloured circles.

In future work, image appearance should be predicted with respect to its colour and spatial surrounding by incorporating the masking effect of colour patterns and measuring the *visibility* of the original contrasts using the contrast sensitivity function (CSF) for chromatic channels [Mullen, 1985] as a function of spatial frequency. Methods for modifying enhancement parameters over time while maintaining temporal coherence should also be investigated. Further work is needed on the problem of converting video to greyscale or to a reduced colour set, as it is important for video stylization, processing and display on limited devices. The combination of global mapping and local chromatic contrast restoration can hopefully be useful for this task, as it has given promising results.

5.6 Summary and Discussion

The problem of converting colour images to greyscale is a classic depiction problem that artists and photographers have struggled with for years. While there are many algorithms for performing the conversion of digital images, most do not guarantee the perceptual accuracy of their results. This section has presented a new approach to colour to grey conversion. The Apparent Greyscale approach offers a more perceptually accurate appearance than standard luminance mapping and generates a closer response to the original than other approaches. In particular, it incorporates the Helmholtz-Kohlrausch effect which is considered fundamental to obtaining faithful greyscale reproductions.

The novel adaptation of unsharp masking adds Cornsweet contours to image regions that are lacking sufficient contrast. The Cornsweet illusion is thus able to create apparent discriminability between regions whose original contrast was mostly chromatic. The two-step treatment is a good compromise between a fully automatic technique (first step) and user control (second step) making this approach well suited for natural images, photographs and artistic reproductions. Finally a major benefit is that consistency among images is preserved by avoiding changes in colour ordering. This makes the technique well adapted to the treatment of videos and image series.

Recently, a perceptual study of colour to greyscale conversion methods by Martin Čadík [2008] found that Apparent Greyscale does indeed produce the most perceptually accurate results, and that it has the most consistent performance on all test images, as summarized by the graph shown in Figure 5.24. The author finds it well suited to colourful images and as producing the most accurate results. In addition, it showed excellent user preference, second only to the contrast enhancing algorithm by Grundland et al. [2007].

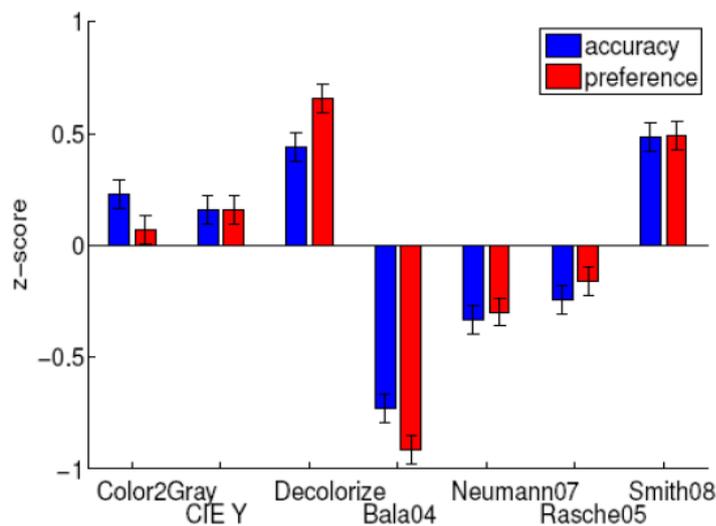


Figure 5.24: Results from perceptual evaluation of greyscale conversion algorithms by [Čadík, 2008]. Apparent Greyscale has the best accuracy and high user preference.

Chapter 6

Enhanced 3D Rendering

A basic principle behind image comprehension is that humans are able to mentally reconstruct an original scenario from visual cues such as shading, occlusions, perspective foreshortening, shadows and specularities. Local contrast enhancement can emphasize these cues, aiding the interpretation of 3D scenes and complex geometry, which is the common task in applications such as medical diagnostics, computer simulations, geographical navigation, game playing and film creation. The main problem is in deciding which cues to emphasize and how to do so with a predictable effect.

To begin, Section 6.1 identifies paintings and artistic techniques that exaggerate local contrasts to alter the visual effect of the painted scene. Section 6.2 gives an overview of computer graphics algorithms developed for enhancing or improving 3D rendered imagery. Section 6.3 presents the novel 3D unsharp masking approach. After explaining how to properly extend 2D unsharp masking to 3D by defining a contrast signal over the 3D mesh, Section 6.4 describes several scenarios to show possible applications.

6.1 In Art: Contours Emphasize Shape and Scene

Local contrast boosting achieved by contour enhancement techniques are common in art, photography and computer graphics. In the creation of photo-real oil paintings, slight non-realistic or exaggerated adjustments are made by the artist. For example, in one of the many still-lives by Willem Claeszoon Heda, the Dutch master who painted incredibly realistic glass and metal objects, one notices how dark and light contours follow along the lemon peel, creating greater separation between it and the table cloth, Figure 6.1. On the left side, a shadow that seems at odds with the lighting, creeps over the lemon's surface to emphasize its shape against the shaded cloth. The peel's twisting shape is accentuated with thin dark contours at regions of self-occlusions. But shape is not the only cue that is enhanced in the painting. Highlights too have a intentional surrounding dark contours, which make the specular appear more pronounced and much brighter than is normally possibly with oil paints on canvas.

In pointillist style, a comprehensive example is Paul Signac's 1948 painting *La salle à manger* shown in Figure 6.2. Luminance contours fall between figures and the background, promoting scene and depth comprehension Figure 6.2(a); dramatic contours along object intersections (like a box lying on a table) increase their salience Figure 6.2(b); and shadows fading towards object edges enhance dark regions where shape

and depth information would otherwise be hidden Figure 6.2(c). Floyd Ratliff suggests that these local adjustments are a natural product of a painter who depicts the contrast he perceives, including physically depicting illusory Mach bands, as mentioned in Section 3.4.



Figure 6.1: Willem Claeszoon Heda masters the depiction of highly reflective surfaces and natural shapes by subtly emphasizing important local contrasts, like the contrasts between the lemon and table cloth, and the contrast between the metallic surface and specular highlight.

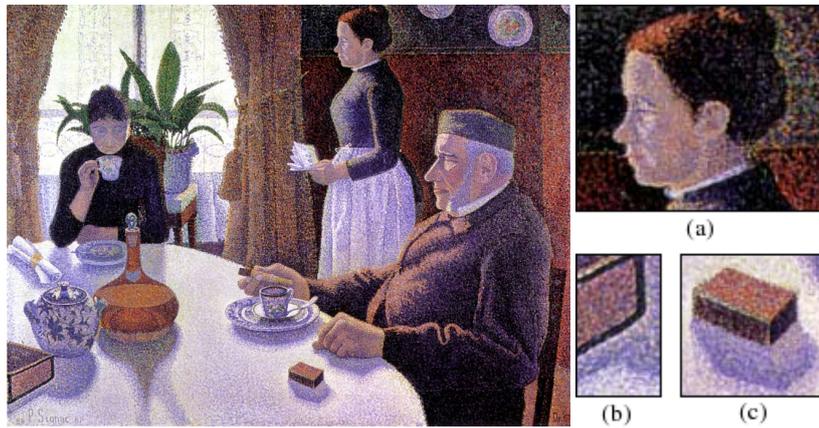


Figure 6.2: Paul Signac's 1948 painting *La salle à manger* is a case of the artist depicting perceived contrast. A bright halo around the woman's head emphasizes figure/ground contrast (a), bright contours in shadowed regions next to dark object edges clarifies the scene's geometry (b) & (c).

6.2 Recent Approaches to Enhance 3D Graphics

Improved depiction of 3D scenes has diverse applications, and as such, has been addressed several times in Computer Graphics. Most recent approaches for creating more evocative renderings perform local enhancement of different aspects, such as incident lighting angle, object features, depth differences. The enhancement of single rendered objects has been of interest for many years, perhaps starting with the seminal paper by Saito and Takahash [1990]. More recently, given a 3D model, Cignoni et al. [2005] shift object normals, performing enhancement to emphasize geometric discontinuities in a single rendered image. In this approach, enhancement strength is controlled by how much the geometry of the mesh is sharpened. The downside is that the shape itself deforms, possibly changing the perception of its form and volume. Also, complex geometry can not be enhanced quickly. Inspired by cartographic illustration, Rusinkiewicz et al. [2006] introduce a new shading model to expose shape features and surface details by positioning a local light per-vertex to achieve maximum contrast. Enhancement strength is controlled by adaptively combining multiple scales of renderings. The resulting images are evocative, however, because of its basis in recalculating the lighting computation, the graphic artist has limited impact and freedom on lighting design.

The enhancement of multi-object scenes addresses the task of communicating the spatial arrangement of various objects and their relationships to each other. To communicate partial self-occlusions of a shape, Ambient Occlusion [Landis, 2002] introduces gradients towards concavities of a surface. In the interest of enhancing visualization of molecules, Tarini et al. [2006] re-shade the renderings to clarify geometry and partial occlusions and draw both gradual wide contours and sharp edges at depth discontinuities, Figure 6.3. In this case, the molecule is not considered a single object, but rather a complex 3D scene in itself. Renderings of medical volumetric data [Bruckner and Gröller, 2007] can benefit from image space halos and enhanced lighting gradients. An aesthetic approach for illustrative game rendering by Mitchell et al. [2007] accentuates internal shape features and adds rim highlights to emphasize object silhouettes.

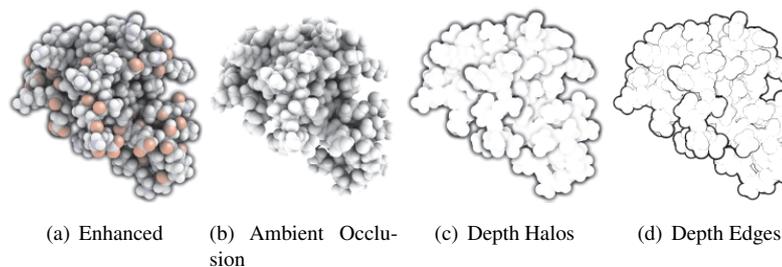


Figure 6.3: Enhanced rendering of a molecule by [Tarini *et al.*, 2006] (a) combines different enhancements, but does not directly enhance 3D cues. Instead, it adds the darkening effects of Ambient Occlusions (b) and contours at depth discontinuities (c) & (d). The depth halos and edges combine to being contours with strong high-frequencies and then lower intensity smooth ramps.

Depth information is key to understanding a multi-object scene. Inspired by artistic techniques, Luft et al. [2006] focus on enhancing depth perception by unsharp masking the depth buffer. They proceed by adapting traditional unsharp masking, described in Section 2.4.4, to the depth buffer by defining the contrast signal as the absolute values of the depth buffer's second derivative. The polarity of change is removed to prevent the

appearance of light contours in deep regions. The contrast signal is scaled according to a user's preference. In the case named "depth darkening", only dark contours are added, which is necessary for giving a strong sense of depth in scenes with long viewing distance. As shown in Figure 6.4(a), the contrast signal produces conflicts between luminance and depth when the more distant side is originally lighter than the near side. Depth unsharp masking falsely lightens the nearer dark regions, producing a blur effect, Figure 6.4(b). A better approach, Figure 6.4(c), would respect the luminance gradient and darken as depth increases when the luminance gradient is very weak by combining depth and luminance information into the contrast signal definition, Figure 6.4(d). In general, the result of depth unsharp masking is that at a depth contrasts, the nearer side is lightened, and the further side is darkened, giving a sense of greater connection to the original 3D information. However, one important cue for spatial understanding is object and ground plane intersections, and since there are no depth discontinuities at these location, the important information is never enhanced.

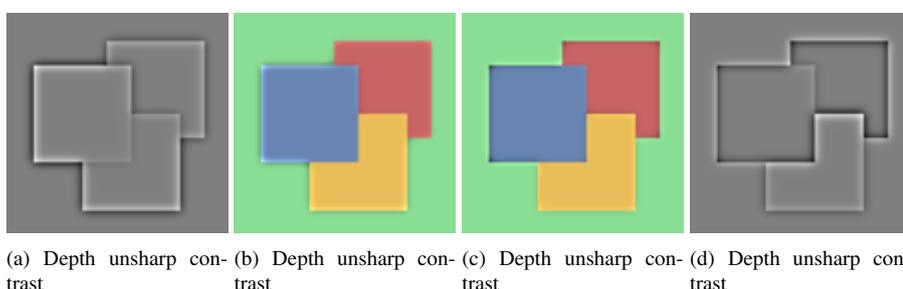


Figure 6.4: Depth unsharp masking defines a contrast signal according to the depth buffer (a) which can cause the lightening and blurred appearance of dark foreground objects (b). A better approach (c) would respect luminance gradient and darken as depth increases when the luminance gradient is very weak by combining depth and luminance information into the contrast signal definition (d).

A trivial extension of depth unsharp masking [Luft *et al.*, 2006] to 3D scenes would define the contrast signal by unsharp masking supplemental 2D buffers like depth (Figure 6.5(a)), object ID (Figure 6.5(b)) and shadows (Figure 6.5(c)). A buffer-based, screen space approach enhances view-dependent discontinuities, like object silhouettes and overlaps. While these are good individual enhancements, when combined, the results appear discontinuous and do not fit within the scene, Figure 6.5(e). In addition, a screen space approach has problems with temporal coherence - imagine the incoherence when occlusions expose/hide very great depth differences. Controllability of the effect is difficult because filter size and enhancement strength must be modified to reflect scene geometry and viewpoint. For some settings, filtering is quite slow and incorporates many pixels into the computation that are not related in the 3D scene. Lastly, for complex objects, the screen space approach is not precise enough because the buffer resolution can not sample the geometry/depth changes at sufficient detail.

Finally, abrupt light changes, like specular highlights or cast shadows, are important for conveying shape, material and spatial arrangements. Highlights in particular ease the perception of shape and material properties [Fleming *et al.*, 2004]. Shadows are necessary cues for proper scene comprehension, and can help by underscoring geometry and clarifying spatial relationships [Cavanagh and Leclerc, 1989]. In recent

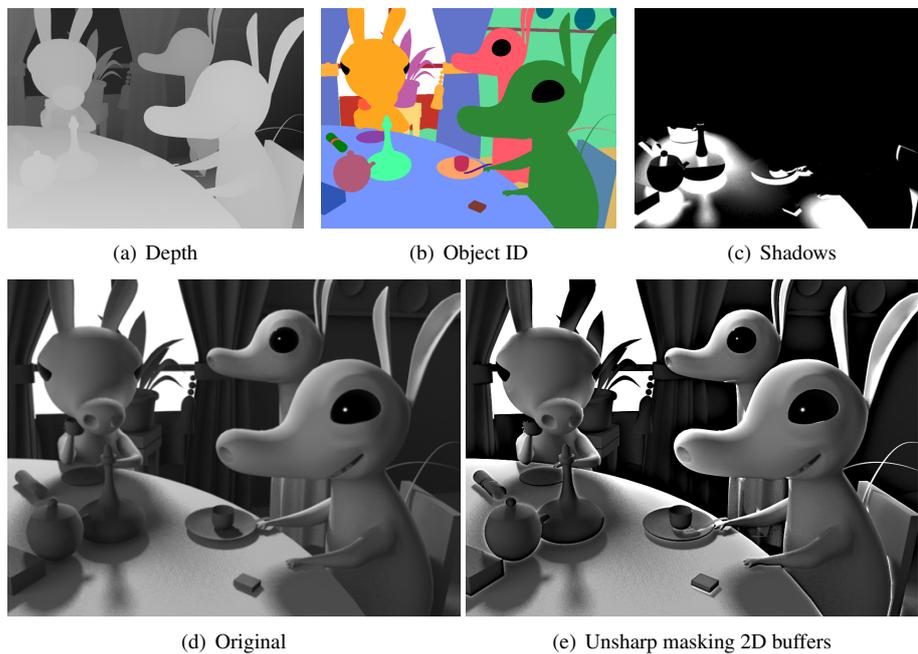


Figure 6.5: Supplemental buffers resulting from rendering a 3D scene (a, b & c). The original rendering (d) and one enhanced by unsharp masking the shadow, contour and depth buffers separately. The normal buffer is not used for this enhancement because there are few fine surface details to enhance. Scene by Saschka Unseld.

work by DeCoro et al. [2007], shadows in image space undergo a recovered perspective transformation, and are then filtered or morphed. Their key contribution is scene coherent abstraction and reshaping of shadows.

6.3 3D Unsharp Masking

This novel work constructs a perceptually founded approach for local contrast enhancement of arbitrary interactive 3D scenes. Rendering has succeeded in producing real to life imagery, and the next step is to make them more communicative and efficient. An enhanced rendering should ease shape recognition, provide better visual separation between objects and clarify their spatial arrangement solely by increasing the apparent contrast of specific visual cues.

The motivations for this endeavor are varied; consider the perusal of a complex anatomical model. To ease comprehension of the complex model, the shading and occlusion gradients should be made obvious, but enhancements that appear or disappear when changing viewpoint would create inconsistencies that impede understanding. A good technique would exaggerate gradients to provide greater visual separation between regions and enhances the shadows that underscore the geometry, doing so coherently over all viewpoints and transitions, and under any type of illumination. Another instance is the exploration of a scene with details that lay in shadow. The desire is for details to remain visible when they are the focus, without forcing the shadows to lighten. By enhancing contrast along the shadow edges, the Cornsweet illusion as de-

scribed in Section 3.2 would create a darker appearance without compromising detail visibility.

Instead of identifying and modifying cues separately, they can be identified by their common cause: spatial variation (gradients) in reflected light. These local contrasts include all cues caused by variations in surface geometry, material properties, incoming light properties and the spatial arrangement of objects. For example, lighting gradients occur where a surface receives different amounts of incoming light (possibly in shadow), where reflectance properties change, and where specular highlights appear. The work presented here simultaneously increases the apparent contrast at all such locations without breaking coherence with the depicted scene. A GPU implementation results in interactive performance and the process fits into the rendering pipeline, allowing simple adjustment of the enhancement strength with an immediately viewable effect. Since this work falls within 3D computer graphics as a last step in the rendering pipeline, its use is not restricted to the creation of a single effective viewing, and can be used in 3D scene exploration and interaction.

6.3.1 Extension from 2D to 3D Unsharp

Unsharp masking in the two-dimensional image domain proceeds first by smoothing the entire image and then adding back the scaled difference between the original and the smooth image, as described in Section 2.4.4. More generally, unsharp masking $\mathcal{U}(S)$ of a signal S , with strength λ and smoothness σ is defined as:

$$\mathcal{U}(S) = S + \lambda(S - S_\sigma).$$

In its 3D adaptation, depicted in Figure 6.6, S is the outgoing radiance from a surface location to a viewpoint, and S_σ is that radiance *smoothed over the surface of the mesh*. For simplicity, the difference between the sharp and smooth signals is denoted as the contrast signal $C(S) = S - S_\sigma$. It is defined over the 3D surface, yet the HDR renderings of signals S and S_σ are used in the per-pixel calculation of the contrast signal $C(S)$. The width of contrast signals are controlled by a user chosen smoothness parameter σ and λ_{Adapt} is a locally adaptive function of a user chosen constant gain λ . By its definition in terms S and S_σ , the highpass signals contained in $C(S)$ lie over the surface, and are thus intrinsically coherent with the scene itself.

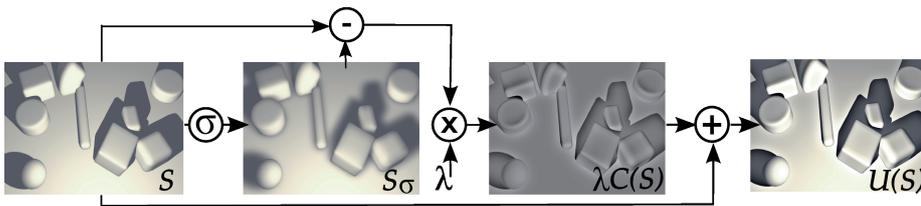


Figure 6.6: The process of 3D unsharp masking sharpens a signal S by adding back a λ -scaled contrast signal $C(S)$, the difference between the original and a smoothed version S_σ .

6.3.2 Preserving Chromatic Saturation Levels

All 3D unsharp masking operations are performed in the CIELAB colour space because it provides a perceptually uniform lightness channel L^* that is decorrelated from its a^*

and b^* chromaticity channels (see Section 2.3.2 for a background on CIELAB). In this way, contrast is measured and modified as apparent lightness and the hue angle $H_{ab} = \tan^{-1}(\frac{b^*}{a^*})$ is unchanged. Formally, the per-channel operations are:

$$\mathcal{U}(S)_{\text{LAB}} = [S_{L^*} + \lambda_{\text{Adapt}} C(S)_{L^*}, kS_{a^*}, kS_{b^*}]$$

Where S_{a^*} and S_{b^*} are scaled by $k = \mathcal{U}(S)_{L^*} / S_{L^*}$ to prevent over- or under-saturation during the lightness adjustment. After enhancement, the new saturation $S_{ab}(\mathcal{U}(S))$ is equal to the original $S_{ab}(S)$, since saturation is defined as $S_{ab} = \sqrt{a^{*2} + b^{*2}} / L^*$.

6.3.3 Adaptive Gain

The gain value is usually a user-chosen constant, however, better results are achieved with a gain value that locally adapts to the scene. With λ , the user balances the amount of enhancement to be applied against overshooting, which appears as visible halos. The halo effect is caused by a deficit of the low frequency signal in S with respect to the high frequencies amplified by adding $C(S)$. The visual system is less sensitive to halos when other frequencies, like those created by a background texture, are present. Thus, in regions where texture masking occurs, λ can be greater without introducing visible halos. To adapt to this perceptual effect, λ is increased in textured regions using a simple model inspired by the quantization strategy in JPEG 2000 [Zeng *et al.*, 2001]. Given a user chosen λ , adaptive gain λ_{Adapt} is defined as:

$$\lambda_{\text{Adapt}} = \lambda \sqrt{1 + \text{Var}(S)} \quad (6.1)$$

where $\text{Var}(S)$ is the variance in the reflected light over a small neighbourhood on the surface of the mesh. Other perceptual issues that affect the appearance of the enhancements are distance from view-point, or perhaps differences in colour. In the future, λ_{Adapt} could be made to account for these aspects.

6.3.4 Contrast Signal Calculation

The calculation of the smooth lighting signal S_{σ} is the integral part of 3D unsharp masking, since it gives rise to the contrast signal $C(S)$. S_{σ} is defined as the view-dependent reflected light S smoothed in 3D over the mesh surface. *Laplacian smoothing* [Taubin, 1995] performed over a triangular mesh surface replaces the intensity of each vertex v_i by the average intensity of its neighbours. Iterating this filter results in increasingly smooth versions of S_{σ} . When σ is small, the contrasts in $C(S) = S - S_{\sigma}$ are narrow (the higher frequencies); as σ increases, the lower frequency bands become included in the contrast signal.

As opposed to 2D smoothing operations, smoothing over a 3D surface automatically adapts to the surface's orientation and location with respect to the viewpoint, thus undergoing correct perspective foreshortening. A hypothetical 2D filter kernel enacting a similar convolution over an image would have to change its 2D orientation, shape and size according to the extracted 3D information. When smoothing is applied directly over the surface, the resulting contrasts are coherent with the scene itself, and the method is often more efficient than a 2D approach attempting to achieve a similar coherence.

The mesh structure and quality affects the results of the enhancements, just as it affects many computer graphics algorithms that operate on meshes. A very low quality

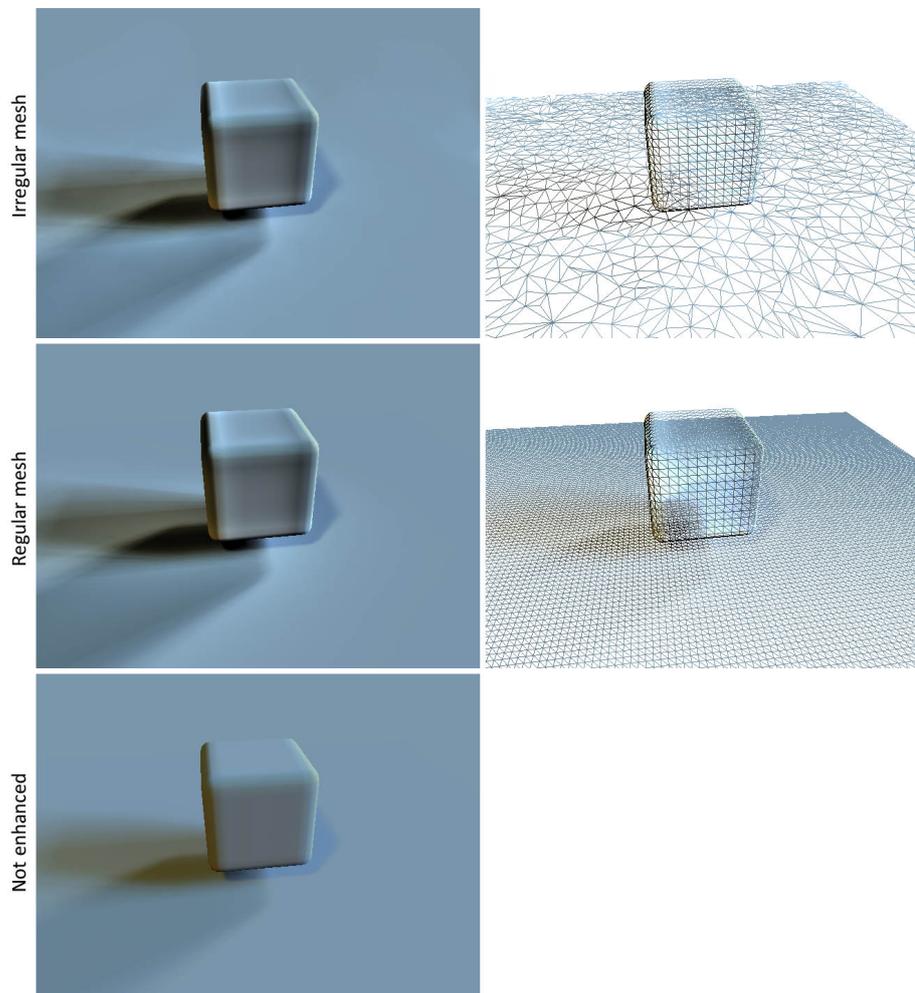


Figure 6.7: When the number of smoothing iterations (σ) is sufficiently high, irregular tessellation does not change the quality of the enhancement. This is because the stronger smoothing blurs the visible meshing artifacts.

mesh will give an irregular smooth signal and less smooth Cornsweet contours. As with most mesh processing algorithms, the calculation of the contrast signal works best when the mesh has uniform vertex density, uniform areas and angles, no isolated triangles or gaps and constant topology. Since the smoothness parameter, σ , depends on the mesh tessellation, the smooth signal S_σ may be more or less evenly/consistently smoothed depending on mesh structure. If the mesh is of lower initial quality and no automatic mesh repair solution is available one can always choose a stronger smoothing (high σ) so that visible meshing will be blurred out, as shown in Figure 6.7.

6.3.5 Comparison of 2D and 3D Unsharp Masking

The enhancement effect of the 3D contrast signal is best understood through a comparison to a 2D unsharp masking contrast signal. In 3D, the signals are continuous along

the surface of the mesh and are not interrupted by distant objects that may be acting as occluders, as would occur in a 2D approach, Figure 6.8(O). This prevents the presence of continuous signals around disconnected objects that change as viewpoint changes. The 3D contrast signal calculation respects gaps between objects so as to be consistent with their spatial grouping. This also prevents brightness changes to a surface on account of differences between it and other distant surfaces, that in 2D lead to brightness inconsistency over time, Figure 6.8(D). In 2D, the detected high-frequencies have the same width in image space, so that near and far gradients are treated in the same scale, when in fact they have different sizes. So a wide gradient that is distant, may be enhanced in the same way as a very narrow gradient, Figure 6.8(P). Lastly, since a 3D calculation of contrast is not affected by what is visible or not visible, temporal popping due to visibility changes is not introduced and there is no spurious enhancement of image borders.

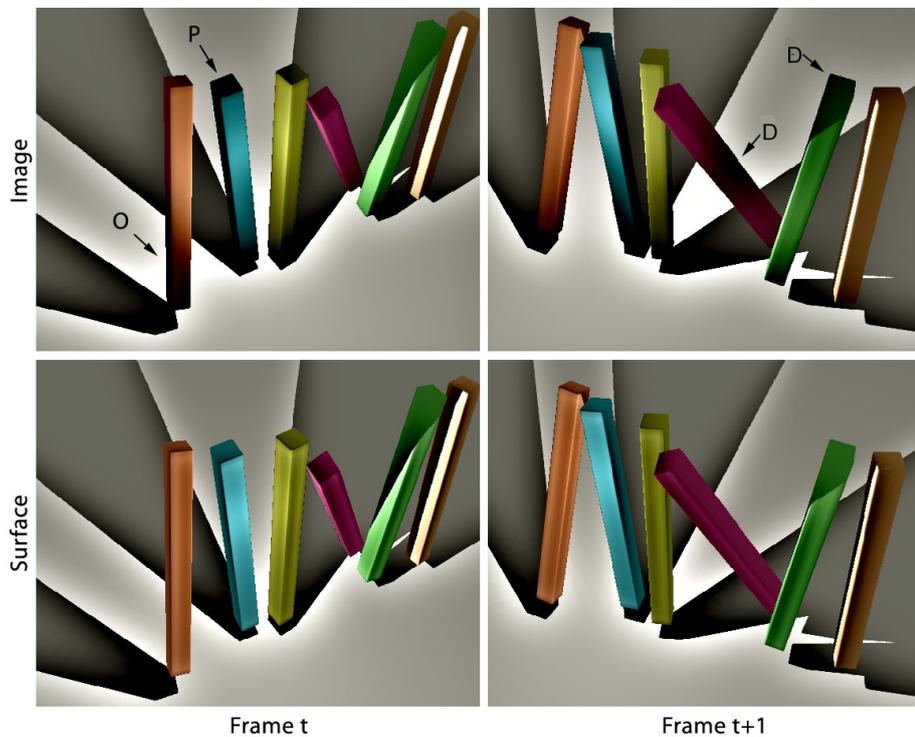


Figure 6.8: Comparing 2D (top) and 3D (bottom) unsharp masking over two successive frames (left and right). **O** (Occlusion, compare vertical): The shadow enhancement should not be interrupted by the occluding stick. **D** (Distant, compare horizontal): Note the strong change in brightness of surfaces between frames because the sticks are darker or lighter than distant background. **P** (Perspective, compare horizontal): Setting a constant filter size that works for shadow edges, disregards discontinuities on the sticks.

6.3.6 Implementation

A main benefit of 3D unsharp masking is its integration in the rendering pipeline and its allowance for interactive parameter adjustment, which is real-time when the scene can be real-time rendered. The 3D unsharp masking algorithm was implemented by Tobias Ritschel with his *Plexus* system for GPU rendering and uses a GPU to perform lighting then smoothing. For implementation details on the smoothing calculation, see [Ritschel *et al.*, 2008]. The main bottleneck is the speed of shadow rendering (256 - 1024 passes), not the smoothing step. This is not the case in the Columns scene, which is most like a game scenario. Its rendering time of 28.3FPS is due to the use of a general shader that works for all lighting scenarios and is not optimized for game scenarios. With an optimized shader, one can anticipate much higher frame rates.

The performance of the approach is dependent on the lighting of each scene. In some cases, this means simple OpenGL point lights with orthographic or cube shadow maps. For natural illumination, the light probe is converted into a number of point lights and a number of shadow maps [Havran *et al.*, 2005]. For the 'Columns' scene, pre-computed ambient occlusion (AO) is used and stored at every vertex, which is common practice in games. Table 6.1 provides a breakdown in terms of lighting computation time, smoothing time, and the remaining computations which comprise the rest, resulting from the implementation using an Nvidia Geforce 8800 GTX, on an Intel Core 2 Duo 6300 1,8 GHz. All tests were done at 640×480 pixel resolution.

| | Scene | Lighting | FPS | | | Vertices | σ | Time | | Supersampling | |
|---|----------|-----------|-------|-------|-------|----------|----------|-------|--------|---------------|-------------|
| | | | Total | W/Out | Extra | | | Light | Smooth | Surface | Framebuffer |
|  | Feet | Natural | 10.2 | 15.2 | 33 % | 57 k | 5 | 26.5 | 3.7 | no | none |
|  | Dice | Point | 15.6 | 63.0 | 75 % | 74 k | 1 | 1.7 | 4.9 | yes | 2×2 |
|  | Keys | Point | 15.2 | 63.0 | 76 % | 152 k | 20 | 5.1 | 34.0 | no | 2×2 |
|  | Columns | Point, AO | 28.3 | 63.2 | 55 % | 119 k | 2 | 7.5 | 2.5 | no | 2×2 |
|  | Chamfer | Natural | 8.3 | 10.7 | 22 % | 39 k | 2 | 20.0 | 10.1 | no | none |
|  | Golfball | Natural | 17.9 | 31.3 | 43 % | 127 k | 8 | 14.3 | 10.3 | no | none |
|  | Cross | Natural | 10.9 | 12.4 | 16 % | 8 k | 10 | 7.2 | 4.7 | no | none |
|  | Lucy | Natural | 9.5 | 37.5 | 75 % | 262 k | 40 | 16.3 | 62.2 | no | none |

Table 6.1: Frame rates for different scenes.

6.3.7 Limitations

For extreme choices of λ , adverse percepts may occur for some viewers. The most common unintended effect is when the enhancement is perceived as a halo. However, this occurs only for extreme choices for λ . Material characteristics may also be enhanced. For instance, shiny objects may appear slightly shinier. This can be interpreted as an advantage that makes cues for material recognition more obvious. 3D unsharp masking technique may enhance artifacts, for instance in soft shadows, by bringing subthreshold errors above visible threshold. To prevent this if it occurs, one could consider separate rendering passes with weaker λ for shadows than meshes. However, this

is not necessary when high quality rendering techniques are used.

It is true that 3D unsharp masking will not explicitly enhance view-dependent cues like object silhouettes. However, it creates enhancements when visible and invisible object sides are differently illuminated (i.e. objects placed on a ground), which gives an effect similar to silhouette enhancement. If explicit enhancement of such cues is desired, then contrast at object boundaries can be increased in one of two ways. Within the 3D unsharp masking framework, placing a light source at the camera will blur the invisible shadow behind the object on the mesh, which creates a contrast signal around the silhouette of the object. Or 3D unsharp masking can simply be complemented by an additional 2D enhancement [Luft *et al.*, 2006; Tarini *et al.*, 2006]. Lastly, the approach will not restore missing contrast, so if a geometric discontinuity is not visible because of the lighting and the object's material, the data is not recovered. It is assumed that lighting does not have a terrible design and is appropriate so that important geometric cues are indeed visible or detectable.

6.4 Applications of 3D Unsharp Masking

The applications of 3D unsharp masking are diverse and include visualization, model analysis, presentation of 3D designs, education, filmmaking and game rendering. Related to the field of perception, 3D unsharp masking could be used to create stimuli for psychophysical experiments on the Cornsweet illusion and other local contrast effects. For other forms of enhancement, instead of being set by a user, λ can be controlled by external 2D or 3D data, such as HDR contrast or lighting intensity. As an NPR tool, the approach could achieve contrast effects, and could be used for enhancing colour, texture or other attributes instead of lightness.

The following set of images are frames rendered with 3D unsharp masking. The general goals of these scenarios are: the enhancement of complex geometry, the extension of a scene's apparent dynamic range, an aid for scene comprehension (i.e. depth, spatial arrangement), coherent artistic effects, the enhancement of a surface's luminance texture and the enhancement of animated and deforming geometry. The temporal coherence of these enhancements are shown in video format online at [Website, 2008a].

To ease comprehension of a complex anatomical model, shading and occlusion gradients are made obvious. This provides greater visual separation between regions and enhances the shadows that underscore the geometry, doing so coherently over all viewpoints and transitions, and under any type of illumination, Figure 6.9. As shown in Figure 6.10, 3D unsharp masking achieves similar high-level enhancements as normals sharpening [Cignoni *et al.*, 2005], exaggerated shading [Rusinkiewicz *et al.*, 2006] and depth unsharp masking [Luft *et al.*, 2006]. Using similar viewpoints and models, the results of 3D unsharp masking are indeed comparable to, or improve over the other methods, and is a comprehensive approach that performs the diverse enhancements with a single simple technique. In addition to making shapes more obvious, depth also appears enhanced due purely to the enhanced shading gradients.

Rusinkiewicz [2006] shows that varying enhancement strength according to importance may be desired. They do so over the space of the rendered image with a multiscale approach. The same can be done within the rendering pipeline, without need for rendering at multiple scales and recombining. This is done by simply painting rough gain controls over the 3D object, in this case to indicate importance with both positive and negative values, and using them as λ . As depicted in Figure 6.11, the effect



Figure 6.9: A rendering of a 3D model may not contain sufficient contrast to emphasize the geometry (leftmost). 3D unsharp masking emphasizes the scene space contrasts including case shadows and self-occlusions under any type of illumination.

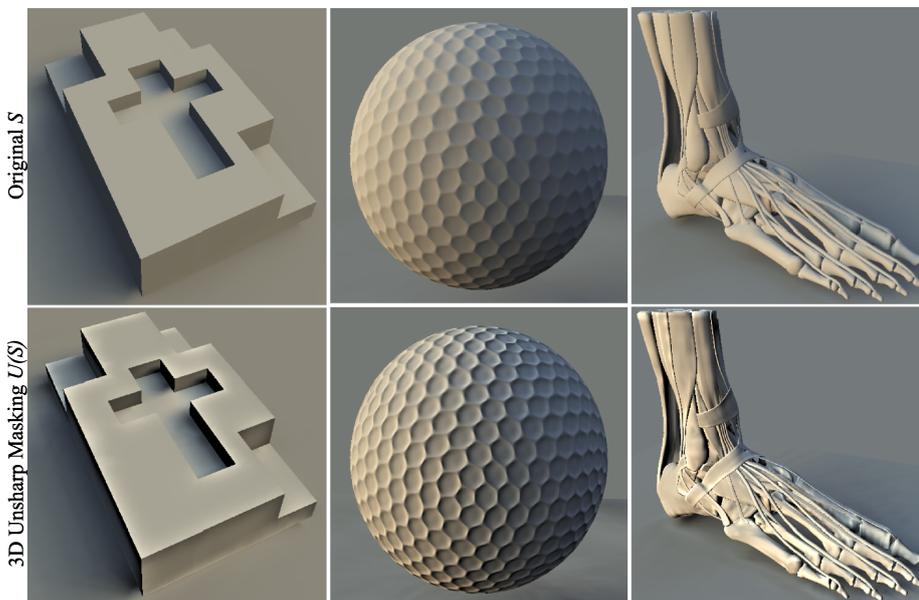


Figure 6.10: Using meshes from Cignoni, Rusinkiewicz and Luft (left to right) under natural lighting, 3D unsharp masking is shown to enhance the renderings in the proper regions and additionally accounts for shadows and contrasts that do not have depth differences.

both enhances the face, torch and hands of the statue while deemphasizing the robe and pedestal.

Returning to the introductory scenario about detail in shadows, 3D unsharp masking successfully creates the impression of darker shadows while keeping the text readable in Figure 6.12. The shadows appear darker because of the Cornsweet illusion, yet do not appear sharpened.

A cluttered scene of similarly shaped keys and screws are shown in Figure 6.13. The resulting enhancements (for two different σ values) ease the cognitive task of object distinction and separation from ground by enhancing the cast shadows, especially those between objects and the ground, and the enhanced speculars aid shape recognition, Figure 6.14. The *Chamfer Plane* scene in Figure 6.15 shows that a naturally lit scene with complex materials and spatial arrangement is seamlessly enhanced. For two different λ values, its soft shadows remain soft, yet slightly more prominent. Occlu-



Figure 6.11: The original rendering (left) is 3D unsharp masked using λ encoded by vertex colour to represent importance, producing imagery (right) that directs attention away from the statue's robe.

sions become more obvious, object silhouettes are emphasized. The glossy highlights become more bright, helping with shape comprehension and curvature understanding. Objects appear more grounded on the plane, since their intersection with the ground is also emphasized.

The same effects created by 3D unsharp masking can be observed in art, where contrasting shades are juxtaposed to exaggerate local lighting discontinuities. For instance, Salvador Dali extends the apparent dynamic range beyond what is physically possible by adding Cornsweet contours along shadow edges, Figure 6.16(a). As in 3D unsharp masking, such contours occur within the context of the depicted scene, not in the context of the 2D painting. Games could opt for approximations of the technique that have neglectable overhead and enhance gradients at strong contrast shadow edges, making environments that look like that in Figure 6.16(c).

Observe that luminance changes created by a texture are also enhanced. As with the book scene shown in Figure 6.12, this simply means the enhancement is less robust to changes in radius and enhancement strength, so users may choose a weaker effect to avoid enhancing the textures. Or, should the user have special reason not to enhance texture at all, he or she may choose to render in two passes (one with reflected lighting,

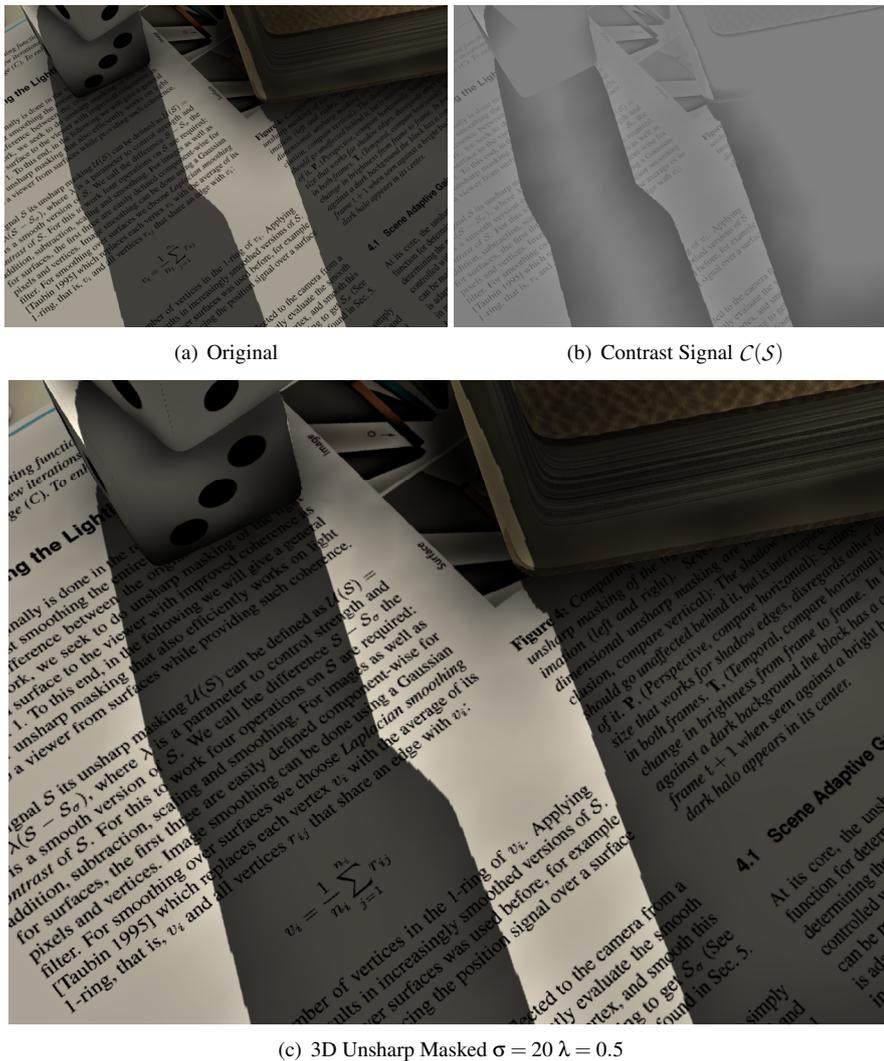


Figure 6.12: The apparent dynamic range of the original (a) is increased by 3D unsharp masking (b), which emphasizes the shadow edges. λ_{Adapt} ensures that the contrast signal is slightly higher at text areas to account for texture masking (c).

one with texture), enhancing only the textureless pass and then recombine (note that this restricts the scene to diffuse textures). Figure 6.17 is an image of a zebra whose stripes are a luminance texture, and its enhancement explicitly shows how 3D unsharp masking seamlessly enhances textures.

Certainly, an advantage of scene-coherent enhancement is that it can be used on animations with deforming models. Two simple examples of 3D unsharp masking applied to animated scenes are presented in the video available at [Website, 2008a]. The first contains only rigid body movement, and the second has smooth deformations. The approach results in high quality enhancements, as shown for selected frames in Figure 6.18. For animated scenes, the direct application of 3D unsharp masking may lead to temporally incoherent enhancements, however, since the lighting function changes

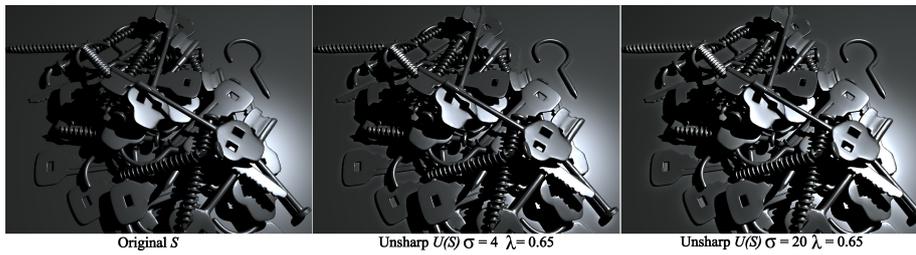


Figure 6.13: 3D unsharp masking enhances cast shadows, emphasizing occlusions. The σ value changes the width of enhancement effects.

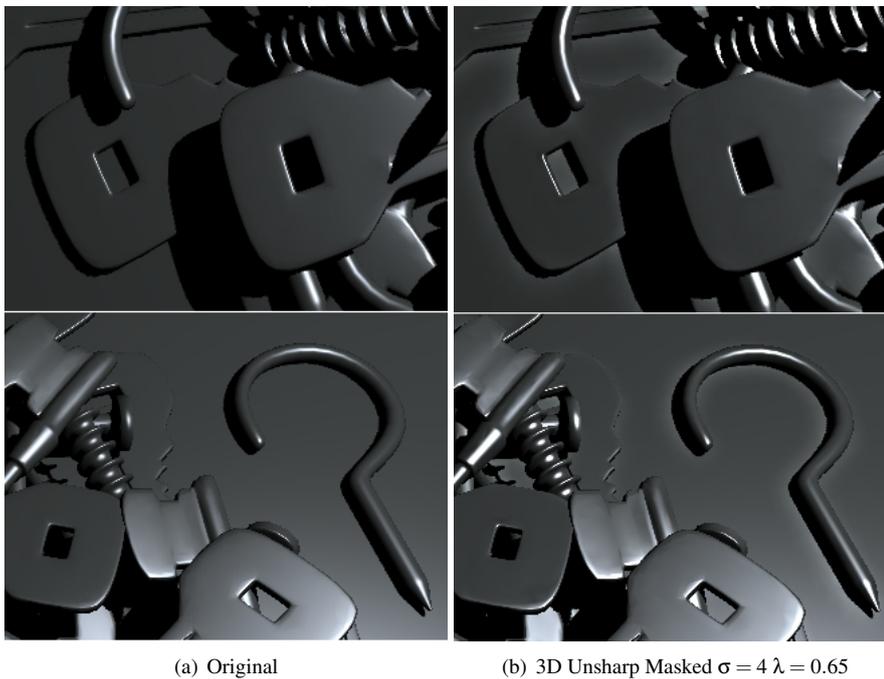


Figure 6.14: Closer inspection shows the enhancing contours at cast shadows and between front- and back-facing regions of the mesh.

smoothly over time, so do the enhancements. The problem of smoothing lighting over the mesh when topology changes or when there are intersections was not investigated.

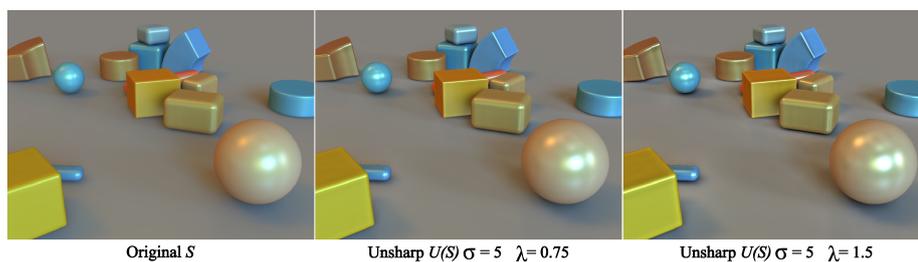


Figure 6.15: 3D unsharp masking coherently enhances soft shadows from natural lighting, object shading changes and specular highlights. Two λ values lead to a subtle or more obvious enhancement; both help to emphasize the objects shapes and their spatial arrangement.

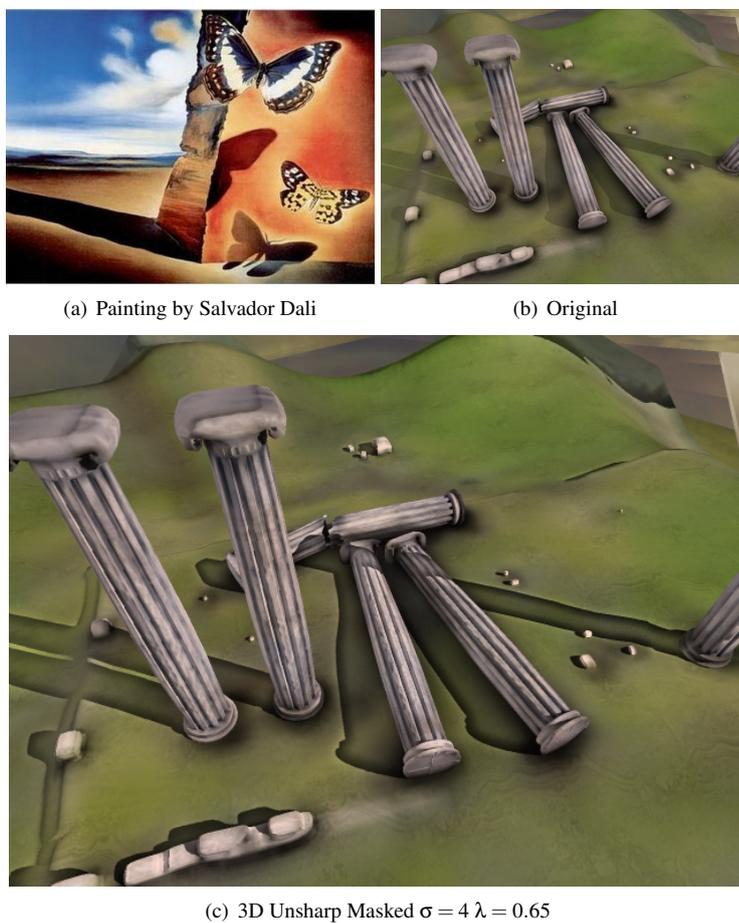


Figure 6.16: Salvador Dali's *Landscape with Butterflies* features visible Cornsweet contours along the shadows that narrow according to perspective. In a stylized game-like environment, the enhancement contours can be visible to emphasize column shape and shadow direction in an artistically inspired style.

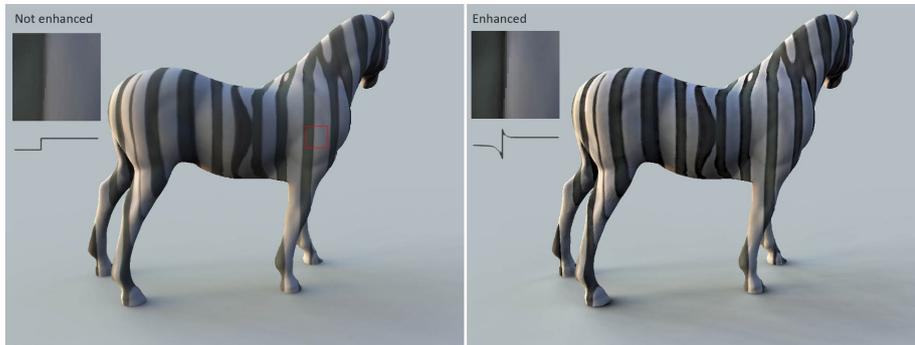


Figure 6.17: The zebra's stripes are given enhanced contrast using the 3D unsharp masking approach.

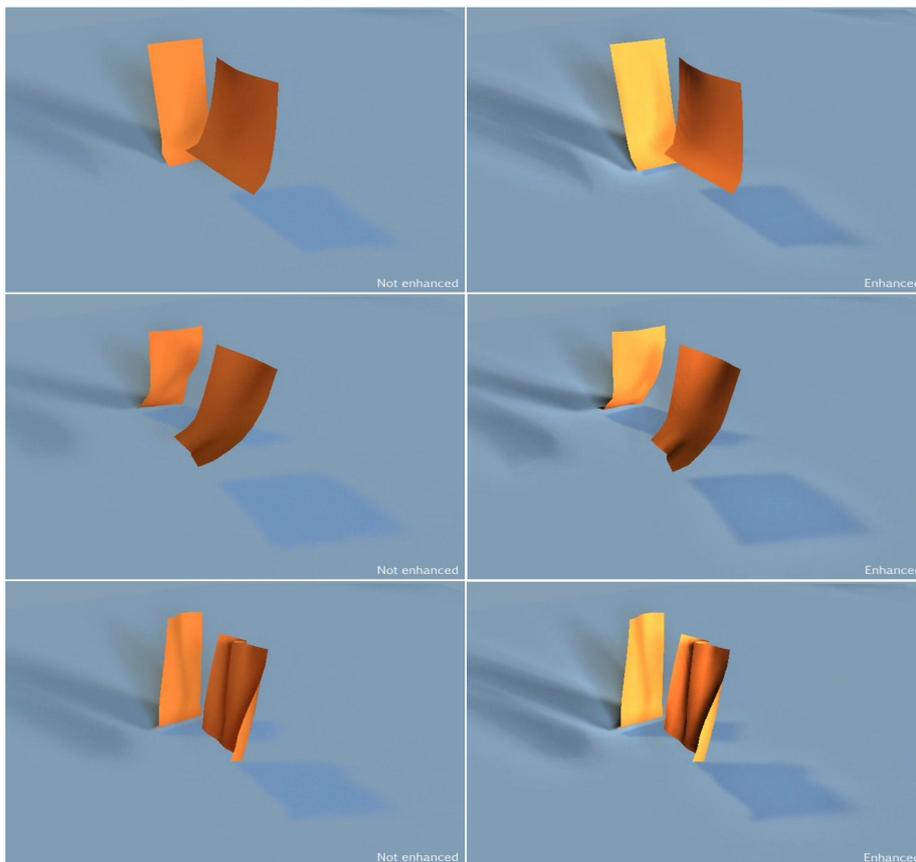


Figure 6.18: Three frames from a simple example of an enhanced animation of a deforming model. In simple cases, without topological changes, temporal coherence is preserved and the 3D unsharp masking can be applied.

6.5 Perceived Effect of 3D Unsharp Masking

Recall the two principles given in the introduction regarding the Cornsweet illusion and the contours that cause it.

First Principle Adding a Cornsweet contour can increase apparent contrast beyond the physical contrast in complex images.

Second Principle Unsharp masking is capable of introducing Cornsweet contours, and the perceptual effect of unsharp masking can be explained by the Cornsweet illusion.

The following two sections give empirical evidence to support both of these claims.

6.5.1 Perceived Contrast Enhancement

To confirm the first principle that apparent contrast is enhanced, results from a perceptual user study proves that 3D unsharp masking increases apparent contrast beyond the original image's contrast. The study measures the effect of parameters and analyzes user reaction to enhancements. Details of the experimental method and extended results from this study have been published in a separate paper by Ihrke et al. [2009]. The major results are presented here. The psychological user study has been designed to confirm that 3D unsharp masking leads to superior and preferred renderings of greater contrast, and that it is easily controlled by the λ parameter. This study does not measure the effect in terms of easier shape or spatial organization understanding, but measures whether a general enhancement is seen and preferred. 15 naïve subjects without prior experience in computer graphics were presented the following diverse scenes: *Chamfer Plane*, *Dice*, *Feet* and *Keys* (Figures 6.15, 6.12, 6.9 and 6.13). For each scene and its set viewpoint, stimuli was generated at three σ values (small, medium and large) over the range $\lambda = [0.0, 3.0]$, sampled at stepsizes below visible threshold. To determine user response, two images were presented side-by-side: the unenhanced image and an adjustable enhanced image controlled by arrow keys to increase or decrease λ . Users were asked to select the 3 images that best responded to the following three questions: first, *When is a difference first visible?* (λ_{Low}), second, *When is the difference too strong or objectionable?* (λ_{High}), and third, *Which image has the best contrast?* (λ_{Best}).

For analysis, each set of λ s were scaled to JND (just-noticeable difference) units based on λ_{Low} . The overall ANOVA test revealed a significant main effect of the three questions ($F(2, 26) = 69.63, p < .001$). Demonstrating the usability of 3D unsharp masking, the λ parameter permits a wide range of acceptable enhancements (≈ 4 JND units on average). These results are depicted in Figure 6.19, which shows the choices of λ_{Low} , λ_{High} and λ_{Best} in JND. Additionally, the σ parameter did not significantly affect the choice of λ_{Best} ($F(2, 28) = 0.69, p = .51$), however, it did impact λ_{High} for two of the scenes, see Figure 6.20. Holm-corrected pairwise statistical contrasts proved that users preferred visible enhancements ($\lambda_{\text{Best}} > \lambda_{\text{Low}}$: $t(14) = 3.06, p < .01$) and that they avoided objectionable artifacts ($\lambda_{\text{Best}} < \lambda_{\text{High}}$: $t = 5.53, p < .001$).

As with nearly all enhancement techniques, the preferred result is subjective. For most of the scenes, users preferred a strong enhancement of approximately twice the visibility threshold (*Feet*: $\lambda_{\text{Best}} = 1.74, p < .01$, *Dice*: $\lambda_{\text{Best}} = 1.97, p < .01$, *Chamfer Plane*: $\lambda_{\text{Best}} = 1.84, p < .01$). For the *Keys* scene, users chose relatively subtle enhancements that were statistically indistinguishable from the lower threshold ($t(14) = 1.28, p = .21$). A cluster-analysis (Ward, euclidean distance) identifies two

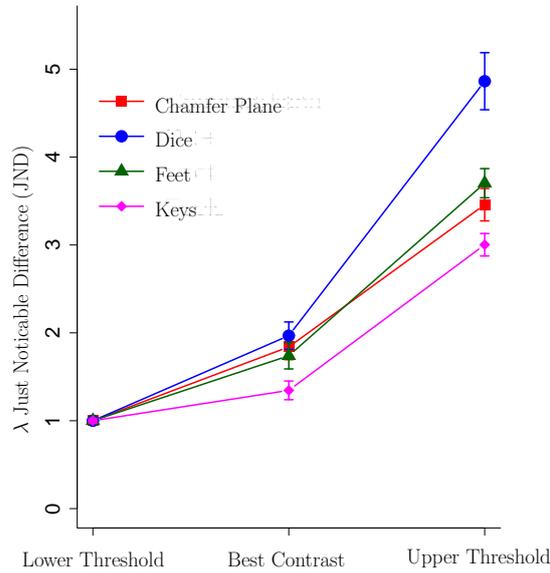


Figure 6.19: User chosen λ 's in JND, averaged over σ for each scene. An enhancement of double λ_{Low} in JND has “best contrast”, tolerable enhancements are up to four times the JND. The error bars are standard errors of the mean (SEM).

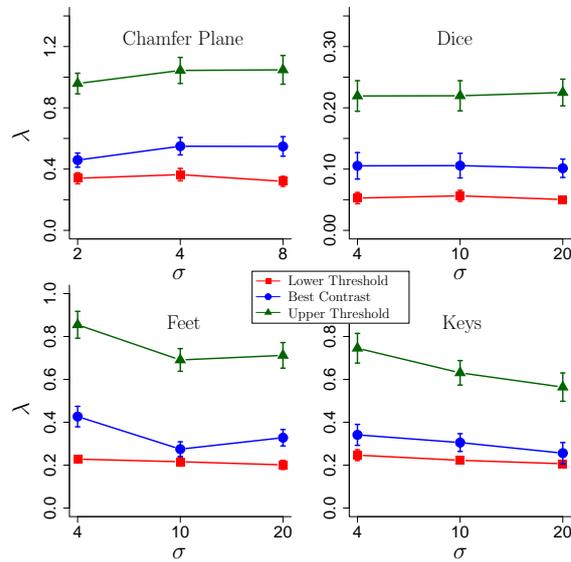


Figure 6.20: The user chosen lower threshold (λ_{Low}), preferred contrast (λ_{Best}), and upper threshold (λ_{High}) as a function of σ . Notice that σ has only a small effect on λ_{Best} .

main groups of users: those who prefer weak enhancement (33%, $\lambda_{\text{Best}} \approx 1.4$ JND) and those who prefer stronger (66%, $\lambda_{\text{Best}} \approx 1.9$ JND). In general, an enhancement of double λ_{Low} in JND is considered to have “best contrast”, and that users tolerate lightness enhancements of up to four times the JND.

6.5.2 3D Cornsweet Illusion

The relationship between unsharp masking and Cornsweet contours was explained in Section 3.4 and there is even stronger evidence linking the Cornsweet illusion and unsharp masking comes arises when moving into 3D. This is unsurprising, since as discussed in Section 3.2.3, Purves et al. [1999] show that the Cornsweet illusion is strongest when coherent with a 3D scene and its lighting. In confirmation of the second principle, 3D unsharp masking does indeed introduce Cornsweet contours. This is proven in Figure 6.21, in which 3D unsharp masking of a uniformly grey cube, lit from below, creates perceived contrast where there is only slight actual contrast.

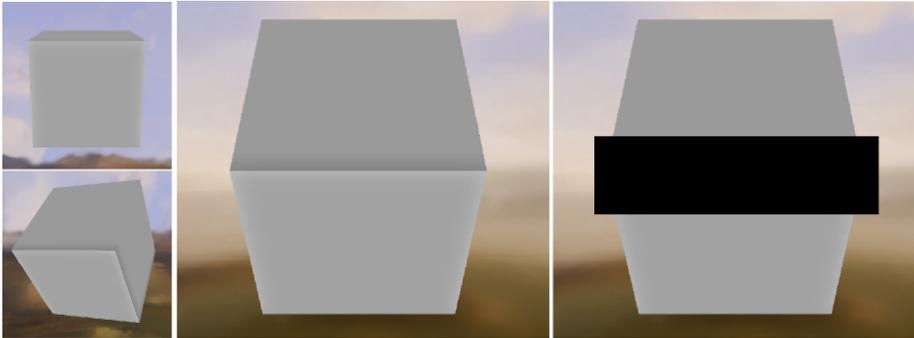


Figure 6.21: Our version of the Cornsweet illusion, made by 3D unsharp masking a uniformly grey cube lit from below.

6.6 Summary and Discussion

Consider the origin of highpass signals contained in $\mathcal{C}(\mathcal{S})$ and their temporal coherence. The signals do not originate from gradients with specific causes such as geometric discontinuities or depth discontinuities. Instead, they include all gradients in reflected light: variations due to curvature and surfaces facing towards or away from light; variations due to texture and highlights; and variations due to different amounts of incoming light. The contrast signal calculation is said to be scene coherent and temporally coherent because it is not affected by view-dependent occlusions, nor is it limited by image space undersampling especially at grazing angles.

The method does not smooth lighting across disconnected components, so it does not properly handle meshes that are not manifold where it is expected. Also, different mesh representations will result in slightly different enhancements. As previously stated, the smooth lighting, described in Section 6.3.4, is sensitive to the mesh tessellation and to the uniformity of length of edges. To avoid uneven smoothing, mesh cleaning operations are applied to all surfaces to refine their tessellation. However, results from the perceptual study (Section 6.5.1) indicate that users do not prefer smaller σ values over larger, so very refined tessellation is not necessary.

Compared to existing approaches, 3D unsharp masking is automatically adapted to the 3D orientation and perspective foreshortening in the scene. Calculating contrast in 3D prevents image space and view-dependent incoherences that are suffered by image space enhancement and using buffered information like depth. Additionally, visually pleasing images are created by adjusting lightness only and correcting colour saturation to prevent the effect of whitening or blackening. Texture masking is taken into account, which allows more contrast to be added where it is less visible. The approach is integrated so that adjustment is simple and responsive, and does not require separate processing.

Chapter 7

Conclusion

This thesis presents ways to depict contrast in images despite constraints imposed by the medium. The motivation is to enhance local contrast so that the resulting imagery communicates the utmost of the represented scene. Three specific situations are addressed: depiction with a limited displayable dynamic range, depiction without the use of colour, and depiction of 3D scenes rendered with fixed lighting and camera exposure. As such, the resulting solutions are new contributions to the fields of tone mapping, image processing and 3D rendering.

When the luminance of a high-dynamic range (HDR) image is compressed to a lower range through tone mapping, its global contrast is reduced and visibility of details lessens. The resulting low-dynamic range (LDR) image may not communicate all the contrasts that were originally present. The luminance of the LDR was carefully adjusted, yet the impression of the depicted scene has become less evocative. A careful treatment of colour can create the appearance of greater global contrast and can restore detail visibility. Better treatment of colour to improve tone mapping results has been suggested, but until now, no method has been realized.

If a greyscale version of a colour image is the desired format, the removal of its chromatic information should not cause its appearance to change. A perceptually accurate conversion from colour to greyscale can preserve all of the original contrast, including differences that were visible purely on account of chromatic information. Many greyscale conversion algorithms exist, from simple global mappings to complicated and expensive optimizations. Yet so far, none have placed perceptual accuracy as the paramount.

Realistic rendering of 3D scenes may not produce the most visually effective depictions. Even with proper lighting and well-chosen view points, shapes may appear ambiguous, spatial organization unclear and texture details less salient. Enhancing its local contrasts can ensure that a greater amount of the scene's visual information is communicated to the viewer, thereby easing the interpretation of complex geometry and the arrangement of objects. Several approaches have focused on enhancing specific visual cues like geometric features and depth discontinuities. However, none have taken a holistic and scene coherent approach to enhancing the appearance of all visual cues necessary for scene comprehension.

7.1 Summary of Contributions

The following novel contrast enhancement algorithms are inspired by the use of local contrast in art and illustration as a way to overcome a variety of depiction constraints. The novel techniques are derived from the standard unsharp masking algorithm for contrast enhancement, which adds high-frequency contours at local contrast locations to create sharper and more vibrant imagery. The success of these novel approaches proves that constraints imposed by the depiction medium can be overcome by modifying traditional unsharp masking into problem-specific solutions.

Enhanced Depiction of Tone Mapped HDR Images

The *Beyond Tone Mapping* method contributes a novel way of adjusting chromatic contrast in an image by unsharp masking the colour channels. For photographers shooting in HDR, it provides a way to incorporate the full captured information in a displayable and printable image. It also means that tone mapping itself is a simpler task, since post-processing can restore diminished or lost contrast. Technically, it contributes a comparison of LDR and HDR contrast, and has wider implications as a luminance-preserving technique for increasing apparent contrast in images.

Perceptually Accurate Images and Video

The *Apparent Greyscale* conversion of colour images and videos to their achromatic counterparts is a multi-scale two-step approach that combines the consistency of a global mapping with the controllability of local enhancement. To address the problem of ordering isoluminant colours, the Helmholtz-Kohlrausch effect is identified and incorporated in the initial mapping from three colour dimensions to one dimensional chromatic lightness. The second step introduces contours at specific locations where the original chromatic contrast has become underrepresented to ensure discrimination of all original colour differences. The algorithm is fast and consistent, so it is the first perceptually accurate algorithm that is well suited to series of images and frame-by-frame greyscale conversion of video. Finally, it does not introduce any spatial changes and preserves the visual impression of the colour version.

Scene Coherent Extension of Unsharp Masking to 3D

3D Unsharp Masking extends the standard 2D technique to a scene coherent enhancement tool that is integrated within the rendering pipeline. The contrasts to be enhanced are defined as intensity variation over the surface of the mesh, and as such, the enhancement of those contrasts avoids the temporal artifacts and 3D inconsistencies that are introduced by 2D image-based approaches. Enhancements can be performed on arbitrarily complex scenes, with arbitrary lighting and the objects can be animated. This holistic approach to enhancing scene renderings is useful for the visualization of models, navigation in virtual environments, stylized game rendering, and creating the appearance of greater dynamic range and depth.

Perceptual

A deeper understanding and justification of the contributed techniques resulted in the identification of a relationship between unsharp masking and the Cornsweet illusion.

First Principle Adding a Cornsweet contour can change the appearance of contrast in complex images.

Second Principle Unsharp masking is capable of introducing Cornsweet contours, and the perceptual effect of unsharp masking can be explained by the Cornsweet illusion.

These claims are empirically justified by observing that unsharp masking adds Cornsweet contours to discriminate between nearly isoluminant colours in greyscale, and by showing that 3D unsharp masking reproduces the 3D Cornsweet illusion. This leads to a final contribution: the presented image enhancement techniques can create stimuli for more complicated perceptual studies into the Cornsweet illusion. *Beyond Tone Mapping* creates chromatic Cornsweet contours, *Apparent Greyscale* creates achromatic Cornsweet contours in complex images that can be controlled at multiple frequency levels, and *3D Unsharp Masking* can create 3D scene coherence Cornsweet contours. Nearly all of the stimuli used for studying the Cornsweet illusion are created by digital artists. With Computer Graphics, the stimuli can be automatically generated over a wide range of parameters, which is a useful contribution to the field of visual perception.

7.2 The Future of Contours in CG

The addition of contours has been shown to be a very powerful tool for enhancing the apparent contrast in depictions, and could have various other applications in computer graphics fields like image processing, visualization and rendering. Furthermore, unsharp masking could be extended to problems in four dimensions. Such an approach could be used in volume rendering to enhance visualizations by smoothing over a lit volume instead of a lit surface.

The existence of contrast models for predicting the effect of Cornsweet contours on apparent contrast could be useful for controlling the amplitude of contrasts added to an image for enhancement. The foremost reason that the perceptual Cornsweet models are not used, is that they are measured on far too simple stimuli. There are no quantitative results measuring the strength of the illusion in a complex image. Other models of sensitivity to high-frequency contrast could also be incorporated as controls, and would better adapt the enhancement to the image.

In the future, the coherence of contours over temporally changing imagery should be addressed. This is true of frame-by-frame 2D enhancements, which could benefit from an incorporation of the previous and following frames. By anticipating the way in which contours animate, temporal coherence could be built in to the enhancement signals. In the 3D case, the coherence of contours should be tested for scenes with deformable meshes, especially those with topological changes and self-intersections. A solution would possibly require an adaption of the 3D unsharp masking smoothing algorithm to the topological changes.

7.3 Concluding Remarks

This thesis champions the use of perceptual models and visual effects in the creation of computer graphics algorithms. A foundation in human vision and perception leads to algorithms that produce more effective imagery, which is a fundamental goal in image

processing and image generation. The depictive quality of an image depends on its appearance, which as shown here, is not always equal to its physically measurable quantities.

Lastly, the quality and usefulness of these contributions have been recognized in the graphics community. The success of the Apparent Greyscale conversion method is confirmed by an independent study, which rates it as the most accurate and widely preferred of all recent approaches [Čadík, 2008]. Based on its consistent and convincing results, an implementation of 3D unsharp masking has been added to a software package for enhancing renderings [Cignoni *et al.*, 2008].

Bibliography

- [Adelson and Ogden, 1984] Anderson C. H. Bergen J. R. Burt P. J. Adelson, E. H. and J. M. Ogden. Pyramid methods in image processing. *RCA Engineer*, 29(6):33–41, 1984.
- [Albers, 1975] Josef Albers. *Interaction of Color*. Yale University Press, New Haven, 1975.
- [Arend, 1994] L. Arend. *Lightness, Brightness, and Transparency*, chapter Intrinsic image models of human color perception, pages 159–213. Hillsdale, NJ: Lawrence Erlbaum Associates, 1994.
- [Badamchizadeh and Aghagolzadeh, 2004] M. A. Badamchizadeh and A. Aghagolzadeh. Comparative study of unsharp masking methods for image enhancement. In *ICIG '04*, pages 27–30, Washington, DC, USA, 2004. IEEE Computer Society.
- [Bala and Braun, 2004] R. Bala and K. Braun. Color-to-grayscale conversion to maintain discriminability. *Proc. SPIE*, 5293:196–202, 2004.
- [Bala and Eschbach, 2004] R. Bala and R. Eschbach. Spatial color-to-grayscale transformation preserving chrominance edge information. *Proc. IS&T/SID's 12th Color Imaging Conference*, pages 82–86, 2004.
- [Barrow and Tenenbaum, 1978] H.G. Barrow and J. Tenenbaum. Recovering intrinsic scene characteristics from images. In *Computer Vision Systems*, pages 3–26. Academic Press, 1978.
- [Bartleson, 1960] C.J. Bartleson. Memory colors of familiar objects. *J. Opt. Soc. Am.*, 50(1):73–77, 1960.
- [Bloj *et al.*, 2007] Marina Bloj, David Connah, and Graham Finlayson. Coding contrast as brightness to convert colour images to greyscale. *J. Vis.*, 7(9):452–452, 6 2007.
- [Braun and Fairchild, 1999] G. Braun and M. Fairchild. Image lightness rescaling using sigmoidal contrast enhancement functions. *Proc. IS&T/SPIE Electronic Imaging 1999*, 1999.
- [Bruckner and Gröller, 2007] Stefan Bruckner and Eduard Gröller. Enhancing depth-perception with flexible volumetric halos. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1344–1351, 2007.

- [Burt and Adelson, 1983] Peter J. Burt and Edward H. Adelson. The Laplacian Pyramid as a compact image code. *IEEE Trans. on Comm.*, pages 532–540, 1983.
- [Čadík, 2008] Martin Čadík. Perceptual evaluation of color-to-grayscale image conversions. *Comput. Graph. Forum*, 27(7), 2008.
- [Calabria and Fairchild, 2003a] A.J. Calabria and M.D. Fairchild. Perceived image contrast and observer preference: Empirical modeling of perceived image contrast and observer preference data. *The Journal of Imaging Science and Technology*, 47:494–508, 2003.
- [Calabria and Fairchild, 2003b] A.J. Calabria and M.D. Fairchild. Perceived image contrast and observer preference: The effects of lightness, chroma, and sharpness manipulations on contrast perception. *J. Imag. Sci. and Tech.*, 47:479–493, 2003.
- [Cavanagh and Leclerc, 1989] P. Cavanagh and YG Leclerc. Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance*, 15(1):3–27, 1989.
- [Cignoni *et al.*, 2005] Paolo Cignoni, Roberto Scopigno, and Marco Tarini. A simple normal enhancement technique for interactive non-photorealistic renderings. *Computer & Graphics*, 29(1), 2005.
- [Cignoni *et al.*, 2008] Paolo Cignoni, Massimiliano Corsini, and Guido Ranzuglia. Meshlab: an open-source 3d mesh processing system, April 2008.
- [Connah *et al.*, 2007] David Connah, Graham Finlayson, and Marina Bloj. Seeing beyond luminance: A psychophysical comparison of techniques for converting colour images to greyscale. In *Proceedings of the 15th Color Imaging Conference*, pages 336–340, 2007.
- [Cornsweet, 1970] T. Cornsweet. *Visual Perception. Chapter II: The Experiment of Hecht, Schlaer, and Pirenne*. Academic Press, 1970.
- [Davey *et al.*, 1998] M.P. Davey, T. Maddess, and M.V. Srinivasan. The spatiotemporal properties of the Craik-O’Brien-Cornsweet effect are consistent with ‘filling-in’. *Vision Research*, 38:2037–2046, 1998.
- [DeCoro *et al.*, 2007] Christopher DeCoro, Forrester Cole, Adam Finkelstein, and Szymon Rusinkiewicz. Stylized shadows. In *International Symposium on Non-Photorealistic Animation and Rendering (NPAR 2007)*, pages 77–83, 2007.
- [deRidder, 1996] H. deRidder. Naturalness and image quality: saturation and lightness variation in color images of natural scenes. *J. Imag. Sci. and Tech.*, 40(6):487–493, 1996.
- [Devinck *et al.*, 2007] Frédéric Devinck, Thorsten Hansen, and Karl R. Gegenfurtner. Temporal properties of the chromatic and achromatic Craik-O’Brien-Cornsweet effect. *Vision research*, 47(27):3385–3393, 2007.
- [Dooley and Greenfield, 1977] R. P. Dooley and Martin I. Greenfield. Measurements of edge-induced visual contrast and a spatial-frequency interaction of the cornsweet illusion. *J. Opt. Soc. Am.*, 67(6), 1977.

- [Durand and Dorsey, 2002] Fredo Durand and Julie Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 257–266, New York, NY, USA, 2002. ACM Press.
- [Durand *et al.*, 2002] Fredo Durand, Maneesh Agrawala, Bruce Gooch, Victoria Interrante, Victor Ostromoukhov, and Denis Zorin. Perceptual and artistic principles for effective computer depiction: Limitations of the medium and pictorial techniques. In *Course Notes #13 (SIGGRAPH 2002)*, 2002.
- [Fairchild and Johnson, 2003] M.D. Fairchild and G.M. Johnson. Image appearance modeling. In *Human Vision and Electronic Imaging VIII*, SPIE, volume 5007, pages 149–160, 2003.
- [Fairchild and Pirrotta, 1991] M.D. Fairchild and E. Pirrotta. Predicting the lightness of chromatic object colors using CIELAB. *Color Research and Application*, 16(6):385–393, 1991.
- [Fairchild, 2005] M.D. Fairchild. *Color Appearance Models, 2nd Ed.* Wiley-IS and T, Chichester, UK, 2005.
- [Fattal *et al.*, 2002] Raanan Fattal, Dani Lischinski, and Michael Werman. Gradient domain high dynamic range compression. *ACM Transactions on Graphics*, 21(3):249–256, July 2002.
- [Fedorovskaya *et al.*, 1997] E.A. Fedorovskaya, H. deRidder, and F.J. Blommaert. Chroma variations and perceived quality of color images of natural scenes. *Color Research and Application*, 22(2):96–110, 1997.
- [Fleming *et al.*, 2004] R. Fleming, A. Torralba, and E. H. Adelson. Specular reflections and the perception of shape. *Journal of Vision*, 4(9):798–820, 2004.
- [Gonzalez and Woods, 2002] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing (2nd Edition)*. Prentice Hall, January 2002.
- [Gooch *et al.*, 2005] Amy A. Gooch, Sven C. Olsen, Jack Tumblin, and Bruce Gooch. Color2gray: salience-preserving color removal. *Proceedings of SIGGRAPH 2005*, 24(3):634–639, 2005.
- [Grundland and Dodgson, 2007] Mark Grundland and Neil A. Dodgson. Decolorize: Fast, contrast enhancing, color to grayscale conversion. *Pattern Recogn.*, 40(11):2891–2896, 2007.
- [Guillon *et al.*, 1998] Sébastien Guillon, Pierre Baylou, Mohamed Najim, and Naamen Keskes. Adaptive nonlinear filters for 2d and 3d image enhancement. *Signal Process.*, 67(3):237–254, 1998.
- [Havran *et al.*, 2005] Vlastimil Havran, Miloslaw Smyk, Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Importance Sampling for Video Environment Maps. In *Eurographics Symposium on Rendering 2005*, pages 31–42,311. ACM SIGGRAPH, 2005.
- [Hunt, 1995] R.W.G. Hunt. *The Reproduction of Color. 5th edition.* Fountain Press, Kingston-upon-Thames, 1995.

- [Ihrke *et al.*, 2009] Matthias Ihrke, Tobias Ritschel, Kaleigh Smith, Thorsten Grosch, Karol Myszkowski, and Hans-Peter Seidel. A perceptual evaluation of 3d unsharp masking. In *To Appear in the 21st Annual IS&T/SPIE Symposium on Electronic Imaging*, 2009.
- [Itten, 1961] Johannes Itten. *The Art of Color*. Van Nostrand Reinhold, New York, 1961.
- [Jobson *et al.*, 1997] Daniel J. Jobson, Ziaur Rahman, and Glenn A. Woodell. A multi-scale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing: Special Issue on Color Processing*, 6(7):965–976, 1997.
- [Kim and Allebach, 2005] S. H. Kim and J. P. Allebach. Optimal unsharp mask for image sharpening and noise removal. *Journal of Electronic Imaging*, 14(2), 2005.
- [Kingdom and Moulden, 1988] F. Kingdom and B. Moulden. Border effects on brightness: A review of findings, models and issues. *Spatial Vision*, 3, 1988.
- [Krawczyk *et al.*, 2007] Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Contrast restoration by adaptive countershading. In *Computer Graphics Forum (Proceedings of Eurographics 2007)*, volume 26, 2007.
- [Kuang *et al.*, 2004] J. Kuang, H. Yamaguchi, G. Johnson, and M. Fairchild. Testing hdr rendering algorithms. In *IS&T/SID Color Imaging Conference*, pages 315–320, 2004.
- [Kuhn *et al.*, 2008] Giovane R. Kuhn, Manuel M. Oliveira, and Leandro A. F. Fernandes. An improved contrast enhancing approach for color-to-grayscale mappings. *Vis. Comput.*, 24(7):505–514, 2008.
- [Landis, 2002] H. Landis. RenderMan in Production. In *ACM SIGGRAPH 2002 Course 16*, 2002.
- [Ledda *et al.*, 2005] Patrick Ledda, Alan Chalmers, Tom Troscianko, and Helge Seetzen. Evaluation of tone mapping operators using a high dynamic range display. *ACM Transactions on Graphics*, 24(3):640–648, 2005.
- [Leu, 2000] J.G. Leu. Edge sharpening through ramp width reduction. 18(6-7):501–514, May 2000.
- [Lin *et al.*, 2006] W.S. Lin, Y.L. Gai, and A.A. Kassim. Perceptual impact of edge sharpness in images. *Vision, Image and Signal Processing, IEE Proceedings -*, 153(2):215–223, April 2006.
- [Lindeberg, 1994] T. Lindeberg. Scale-space theory: A basic tool for analysing structures at different scales. *J. of Applied Statistics*, 21(2):224–270, 1994.
- [Livingstone, 2002] Margaret Livingstone. *Vision and Art: The Biology of Seeing*. Harry N. Abrams, 2002.
- [Luft *et al.*, 2006] Thomas Luft, Carsten Colditz, and Oliver Deussen. Image enhancement by unsharp masking the depth buffer. *ACM Transactions on Graphics*, 2006. To Appear.

- [Majumder and Irani, 2007] Aditi Majumder and Sandy Irani. Perception-based contrast enhancement of images. *ACM Trans. Appl. Percept.*, 4(3):17, 2007.
- [Mantiuk *et al.*, 2006a] Rafal Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Applied Perception*, 3(3), 2006.
- [Mantiuk *et al.*, 2006b] Rafał Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. A perceptual framework for contrast processing of high dynamic range images. *ACM Transactions on Applied Perception*, 3(3):286–308, July 2006.
- [Matkovic *et al.*, 2005] Kresimir Matkovic, Laszlo Neumann, Attila Neumann, Thomas Psik, and Werner Purgathofer. Global contrast factor—a new approach to image contrast. In *Computational Aesthetics in Graphics, Visualization and Imaging 2005*, pages 159–168, 2005.
- [Mitchell *et al.*, 2007] Jason Mitchell, Moby Francke, and Dhabih Eng. Illustrative rendering in team fortress 2. In *NPAR '07: Proceedings of the 5th international symposium on Non-photorealistic animation and rendering*, pages 71–76, New York, NY, USA, 2007. ACM.
- [Mullen, 1985] K. T. Mullen. The contrast sensitivity of human color vision to red-green and blue-yellow chromatic gratings. *Journal of Psychology*, 359:381–400, 1985.
- [Nayatani, 1997] Y. Nayatani. Simple estimation methods for the Helmholtz-Kohlrausch effect. *Color Res. Appl.*, 22:385–401, 1997.
- [Nayatani, 1998] Y. Nayatani. Relations between the two kinds of representation methods in the Helmholtz-Kohlrausch effect. *Color Res. Appl.*, 23:288–301, 1998.
- [Neumann *et al.*, 2007] Laszlo Neumann, Martin Cadik, and Antal Nemcsics. An efficient perception-based adaptive color to gray transformation. In *Proc. of Computational Aesthetics 2007*, pages 73–80, 2007.
- [Nowak and Baraniuk, 1998] Robert D. Nowak and Richard G. Baraniuk. Adaptive weighted highpass filters using multiscale analysis. *IEEE Transactions on Image Processing*, 7(7):1068 – 1074, 1998.
- [O’Brien, 1958] Vivian O’Brien. Contour perception, illusion, and reality. *J. opt. Soc. Amer.*, 48:112–119, 1958.
- [Palmer, 1999] S.E. Palmer. *Vision Science: Photons to Phenomenology*, chapter 3.3 Surface-Based Color Processing. The MIT Press, 1999.
- [Pattanaik *et al.*, 1998] Sumanta N. Pattanaik, James A. Ferwerda, Mark D. Fairchild, and Donald P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. *Computer Graphics*, 32(Annual Conference Series):287–298, 1998.
- [Pattanaik *et al.*, 2000] S.N. Pattanaik, J.E. Tumblin, H Yee, and D.P. Greenberg. Time-dependent visual adaptation for realistic image display. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, pages 47–54, July 2000.

- [Peli, 2000] E. Peli. Contrast in complex images. *J. Opt. Soc. Am. A*, 7(10):2032–2039, 2000.
- [pfsTMO, 2008] pfsTMO. Open source implementation of tone mapping operators, 2008. <http://www.mpii.mpg.de/resources/tmo/>.
- [Polesel *et al.*, 2000] A. Polesel, G. Ramponi, and V. Mathews. Image enhancement via adaptive unsharp masking. *IEEE Trans. Image Processing*, 9:505–510, 2000.
- [Purves *et al.*, 1999] D. Purves, A. Shimpi, and R. B. Lotto. An empirical explanation of the cornsweet effect. *J. of Neurosci*, 19:8542–8551, 1999.
- [Ramponi, 1998] G. Ramponi. A cubic unsharp masking technique for contrast enhancement. *Signal Processing*, 67(2):211–222, June 1998.
- [Rasche *et al.*, 2005] K. Rasche, R. Geist, and J. Westall. Re-coloring images for gamuts of lower dimension. *Computer Graphics Forum, proceedings of Eurographics 2005*, 24(3), 2005.
- [Ratliff, 1965] Floyd Ratliff. *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. Holden-Day, Inc., San Francisco, CA, USA, 1965.
- [Ratliff, 1971] Floyd Ratliff. Contour and contrast. *Proceedings of the American Philosophical Society*, 115(2):150–163, 1971.
- [Ratliff, 1996] Floyd Ratliff. The theory of color and the practice of painting. *Color Function Painting: The Art of Josef Albers, Julian Stanczak and Richard Anuszkiewicz*, pages 8–15, 1996.
- [Reinhard and Devlin, 2005] E. Reinhard and K. Devlin. Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Visualization and Computer Graphics*, 11(1):13–24, 2005.
- [Reinhard *et al.*, 2001] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley. Color transfer between images. *IEEE Comput. Graph. Appl.*, 21(5):34–41, 2001.
- [Reinhard *et al.*, 2002] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. In *ACM Transactions on Graphics*, pages 267–276, New York, NY, USA, 2002. ACM Press.
- [Reinhard *et al.*, 2005] Erik Reinhard, Greg Ward, Sumanta Pattanaik, and Paul Debevec. *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*. Morgan Kauffman, 2005.
- [Ritschel *et al.*, 2008] Tobias Ritschel, Kaleigh Smith, Matthias Ihrke, Thorsten Grosch, Karol Myszkowski, and Hans-Peter Seidel. 3d unsharp masking for scene coherent enhancement. *ACM Transactions on Graphics*, 27(3), 2008.
- [Rusinkiewicz *et al.*, 2006] Szymon Rusinkiewicz, Michael Burns, and Doug DeCarlo. Exaggerated shading for depicting shape and detail. *ACM Trans. Graph SIGGRAPH*, 25(3), July 2006.
- [Saito and Takahash, 1990] T. Saito and T. Takahash. Comprehensible rendering of 3-d shapes. *Computer Graphics*, 24(4):197–206, 1990.

- [Schreiber, 1970] William F. Schreiber. Wirephoto quality improvement by unsharp masking. *Pattern Recognition*, 2:117–121, 1970.
- [Scognamiglio *et al.*, 1999] G. Scognamiglio, G. Ramponi, A. Rizzi, and L. Albani. A rational unsharp masking method for tv applications. *International Conference on Image Processing*, 4:247–251, 1999.
- [Smith *et al.*, 2006] Kaleigh Smith, Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Beyond tone mapping: Enhanced depiction of tone mapped HDR images. In *The European Association for Computer Graphics 27th Annual Conference: EUROGRAPHICS 2006*, volume 25(3), pages 427–438, 2006.
- [Smith *et al.*, 2008] Kaleigh Smith, Pierre-Edouard Landes, Joëlle Thollot, and Karol Myszkowski. Apparent greyscale: A simple and fast conversion to perceptually accurate images and video. *Computer Graphics Forum (Proceedings of Eurographics 2008)*, 27(2), 2008.
- [Spencer *et al.*, 1995] Greg Spencer, Peter S. Shirley, Kurt Zimmerman, and Donald P. Greenberg. Physically-based glare effects for digital images. In *Proceedings of SIGGRAPH 95*, Computer Graphics Proceedings, Annual Conference Series, pages 325–334, August 1995.
- [Stevens and Stevens, 1960] S.S. Stevens and J.C. Stevens. Brightness function: parametric effects of adaptation and contrast. *J. Optical Soc. of America*, 50(11):1139A, November 1960.
- [Strickland *et al.*, 1987] R. N. Strickland, C.-S. Kim, and W. F. McDonnell. Digital color image enhancement based on the saturation component. *Optical Engineering*, 26:609–616, jul 1987.
- [Tarini *et al.*, 2006] Marco Tarini, Paolo Cignoni, and Claudio Montani. Ambient occlusion and edge cueing for enhancing real time molecular visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1237–1244, 2006.
- [Taubin, 1995] Gabriel Taubin. A signal processing approach to fair surface design. In *ACM Trans. Graph. SIGGRAPH*, pages 351–358, 1995.
- [Thomas and Rodriguez, 1997] R.N. Thomas, B.A. and Strickland and J.J. Rodriguez. Color image enhancement using spatially adaptive saturation feedback. *Proceedings International Conference on Image Processing*, 3, 1997.
- [Tumblin and Rushmeier, 1993] Jack Tumblin and Holly E. Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics and Applications*, 13(6):42–48, November 1993.
- [Tumblin and Turk, 1999] Jack Tumblin and Greg Turk. Lcis: a boundary hierarchy for detail-preserving contrast reduction. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 83–90, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.
- [Tumblin *et al.*, 1999] Jack Tumblin, Jessica K. Hodgins, and Brian K. Guenter. Two methods for display of high contrast images. *ACM Transactions on Graphics*, 18(1):56–94, January 1999. ISSN 0730-0301.

- [Tumblin, 1999] J.E. Tumblin. *Three Methods For Detail-Preserving Contrast Reduction For Displayed Images*. PhD thesis, Georgia Institute of Technology, 1999.
- [Valois and Valois, 1990] Russell L. De Valois and Karen K De Valois. *Spatial Vision*. Oxford University Press, 1990.
- [Wachtler and Wehrhahn, 1997] T. Wachtler and C. Wehrhahn. The Craik-O'Brien-Cornsweet illusion in colour: Quantitative characterisation and comparison with luminance. *Perception*, 26:1423–1430, 1997.
- [Wang *et al.*, 2001] Y. Wang, Q. Wu, K. Castleman, and Z. Xiong. Image enhancement using multiscale differential operators, 2001.
- [Ward Larson *et al.*, 1997] G. Ward Larson, H. Rushmeier, and C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Trans. Vis. Comput. Graph.*, 3(4):291–306, 1997.
- [Ware and Cowan, 1983] C. Ware and W.B. Cowan. The chromatic Cornsweet effect. *Vision Research*, 23:1075–1077, 1983.
- [Website, 2008a] 3D Unsharp Masking Website. 3D Unsharp Masking project <http://www.mpi-inf.mpg.de/resources/3DUnsharpMasking/>, 2008.
- [Website, 2008b] Apparent Greyscale Website. Apparent Greyscale Project <http://www.mpi-inf.mpg.de/resources/ApparentGreyscale/>, 2008.
- [Winkler, 2005] S. Winkler. *Digital Video Quality: Vision Models and Metrics*. John Wiley & Sons, Ltd, West Sussex, England, 2005.
- [Wyszecki, 1967] G. Wyszecki. Correlate for lightness in terms of CIE chromaticity coordinates and luminous reflectance. *Journal of the Optical Society of America (1917-1983)*, 57, February 1967.
- [Yoshida *et al.*, 2005] Akiko Yoshida, Volker Blanz, Karol Myszkowski, and Hans-Peter Seidel. Perceptual evaluation of tone mapping operators with real-world scenes. In *Human Vision and Electronic Imaging*, SPIE, volume 5666, pages 192–203, 2005.
- [Zeng *et al.*, 2001] Wenjun Zeng, Scott Daly, and Shawmin Lei. An overview of the visual optimization tools in JPEG 2000. *Signal Processing: Image communication Journal*, 17(1):85–104, 2001.