

Geometric Methods in Perceptual Image Processing

A dissertation presented

by

Hamilton Yu-Ik Chong

to

The School of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Computer Science

Harvard University

Cambridge, Massachusetts

May 2008

©2008 - Hamilton Yu-Ik Chong

All rights reserved.

Thesis advisor

Author

Steven J. Gortler

Hamilton Yu-Ik Chong

Geometric Methods in Perceptual Image Processing

Abstract

Perceptual image processing refers to the algorithmic transformation of information in which input images are turned into inferred descriptions (e.g., three-dimensional shape and material properties) of the objects being viewed. The ability of humans to derive such information—even when such problems are ill-posed—demonstrates the utility of modeling aspects of human visual perception for use in graphics and vision applications.

In this thesis, we first study the problem of color constancy and characterize the applicability of the generalized von Kries models in terms of rank constraints on the measured world. We show that our world, as measured by some spectral databases, approximately meets these rank constraints and we provide an algorithm for computing an optimal color basis for generalized von Kries modeling. These color constancy notions are then used to derive a new color space for illumination-invariant image processing (in which algorithms manipulate the intrinsic image instead of working directly on RGB values). The derived color space also possesses other useful perceptual features: Euclidean distances approximate perceptual distances, and the coordinate directions have an intuitive interpretation in terms of color opponent channels.

Finally, we draw some connections between curves in an image and shape un-

derstanding. We single out suggestive contours and illumination valleys as particularly interesting because although one is defined in terms of three-dimensional geometry and the other in terms of image features, the two produce strikingly similar results (and effectively convey a sense of shape). This suggests that the two types of curves capture similar pieces of geometric information. To explore this connection, we develop some general techniques for recasting questions about the image as questions about the surface.

Contents

| | |
|--|-----------|
| Title Page | i |
| Abstract | iii |
| Table of Contents | v |
| Citations to Previously Published Work | viii |
| Acknowledgments | ix |
| Dedication | xi |
| 1 Introduction | 1 |
| 1.1 Perceptual Image Processing | 1 |
| 1.1.1 Color Constancy | 3 |
| 1.1.2 A Perception-based Color Space | 7 |
| 1.1.3 Shape Perception and Line Drawings | 9 |
| 1.2 Contributions | 11 |
| 1.3 Overview of Thesis | 13 |
| 2 A Basis for Color Constancy | 16 |
| 2.1 Introduction and Previous Work | 16 |
| 2.2 Theory | 19 |
| 2.2.1 Measurement Constraints | 22 |
| 2.3 Color Basis for Color Constancy | 25 |
| 2.4 Relationship to Previous Characterizations | 27 |
| 2.5 Results | 30 |
| 2.5.1 Effective Rank of “the World” | 30 |
| 2.5.2 Von Kries Sensors | 31 |
| 2.5.3 White Patch Normalization | 33 |
| 2.5.4 White Balancing | 38 |
| 2.5.5 Constrained Optimization | 40 |
| 2.6 Discussion | 43 |
| 3 A New Color Space for Image Processing | 44 |
| 3.1 Introduction | 45 |

| | | |
|----------|--|------------|
| 3.2 | Previous work | 48 |
| 3.3 | A Perceptual Metric | 50 |
| 3.3.1 | Formalizing the Color Space Conditions | 50 |
| 3.3.2 | Form of the Metric | 51 |
| 3.3.3 | Parameter Estimation | 52 |
| 3.4 | Metric Results | 56 |
| 3.5 | Applications | 57 |
| 3.5.1 | Segmentation | 58 |
| 3.5.2 | Poisson Image Editing | 60 |
| 3.6 | Discussion | 65 |
| 4 | Differential Geometry | 67 |
| 4.1 | Manifolds, Tensors, and Calculus | 68 |
| 4.1.1 | Manifolds | 69 |
| 4.1.2 | Tensor Fields | 70 |
| 4.1.3 | Integrals | 79 |
| 4.1.4 | Derivatives | 83 |
| 4.2 | Intrinsic and Extrinsic Properties of Surfaces | 89 |
| 4.2.1 | First Fundamental Form | 89 |
| 4.2.2 | Second Fundamental Form | 90 |
| 4.2.3 | Tensors that live on the surface itself | 92 |
| 4.2.4 | Higher Order Derivatives | 94 |
| 4.3 | The Metric Equivalence Problem | 96 |
| 5 | Shapes from Curves | 98 |
| 5.1 | Introduction | 99 |
| 5.2 | Curve Definitions | 101 |
| 5.2.1 | Surface-only Curves | 102 |
| 5.2.2 | Environment-dependent Curves | 106 |
| 5.3 | Relations Between Image and Surface Curves | 116 |
| 5.3.1 | Image Plane and Orthographic Projection | 116 |
| 5.3.2 | Critical Points of Illumination | 124 |
| 5.3.3 | Saint-Venant and Suggestive Energies | 126 |
| 5.3.4 | Suggestive Contours and Shading | 128 |
| 5.4 | Discussion | 129 |
| 6 | Conclusions and Future Work | 130 |
| | Bibliography | 133 |
| A | Color Constancy Proofs | 138 |
| A.1 | Conditions for Success | 138 |

| | | |
|----------|--|------------|
| A.1.1 | Proof of Proposition 1 | 138 |
| A.1.2 | Proof of Lemma 1 | 141 |
| A.2 | The Space of Solutions | 145 |
| B | Color Space Proofs | 147 |
| B.1 | Deriving the Functional Form | 147 |
| B.2 | Recovering Weber’s Law | 149 |
| C | More Differential Geometry | 150 |
| C.1 | Non-coordinate Bases | 150 |
| C.2 | Structure Equations for Surfaces | 154 |
| C.3 | Theorema Egregium | 159 |
| D | Image and Surface Curves | 160 |
| D.1 | Basic Facts | 160 |
| D.2 | Screen Coordinate Vector Fields | 166 |
| D.3 | Curve Relations | 168 |
| D.4 | Principal Directions | 175 |
| D.5 | Apparent Curvature | 176 |
| D.6 | General Lighting | 180 |

Citations to Previously Published Work

Large portions of Chapter 2 on color constancy have previously appeared in:

“The von Kries Hypothesis and a Basis for Color Constancy”.

H. Y. Chong, S. J. Gortler, T. Zickler.

In *Proceedings of ICCV 2007*.

The perceptual optimization of a basis for color constancy (Section 2.5.5) and the development of a new color space as presented in Chapter 3 have previously appeared in:

“A Perception-based Color Space for Illumination-invariant Image Processing”.

H. Y. Chong, S. J. Gortler, T. Zickler.

In *Proceedings of SIGGRAPH 2008*.

Acknowledgments

I would first like to thank my advisor Professor Steven Gortler for his support over the years (during both undergraduate and graduate days). From the beginning, he gave me great freedom in choosing topics to pursue and provided excellent guidance on how to approach any chosen problem. I would also like to thank Professors Todd Zickler, Fredo Durand, Roger Brockett, and Craig Gotsman, all of whom have served on at least one of my oral exam committees (and read my various scribbles). They have served as wonderful sounding boards for my (often outlandish) ideas and thoughts. Thanks also to Professors Doug DeCarlo and Szymon Rusinkiewicz for very helpful discussions on curves and surfaces. And thanks as well to Brian Guenter for his mentorship at Microsoft Research and for giving me the opportunity to broaden my research exposure and interact with the amazing full-timers, postdocs, and fellow interns there.

Graduate school is of course a sometimes trying experience, and without the mutual support of friends—all going through similar trials and tribulations—the hurdles would likely feel quite insurmountable. So special thanks to Michael Kellermann, Yi-Ting Huang, Jimmy Lin, Daniel DeSousa, Eleanor Hubbard, Leo Nguyen, Michelle Gardner, James Black, Mark Hempstead, and Camilo Libedinsky, who put up with my antics outside of lab and joined me for sports and gaming. Many thanks as well to lab-mates Danil Kirsanov, Guillermo Diez-Cañas, Brenda Ng, Loizos Michael, Philip Hendrix, Ece Kamar, Geetika Lakshmanan, Doug Nachand, Yuriy Vasilyev, Fabiano Romiero, Emmanuel Turquin, Christopher Thorpe, Charles McBrearty, Zak Stone, Kevin Dale, Kalyan Sunkavalli, Moritz Baecher, Miriah Meyer, Christian Ledergerber, and Forrester Cole, who made the lab an inviting place. Thanks also to David Harvey, Wei Ho, and Ivan Petrakiev for their

math pointers (and conversations outside of math as well). Thanks also to my friends in and around Saratoga.

And of course, profuse thanks goes to my family for their constant encouragement (despite their teasing me about the “cold” California winters). I certainly would not have gotten far at all without their support. I am reminded of a Chinese adage (popularized by Yao Ming I believe): “How do I thank my family? How does a blade of grass thank the sun?” Well, I do not have an answer to that question, so I will have to leave it at just “Thanks!”

Thanks again to everyone (and apologies to the many other deserving people I’ve left unmentioned)!

*Dedicated to my parents Fu-Chiung and Kathleen Chong,
to my brothers Sanders and Anthony Chong,
and to all teachers.*

Chapter 1

Introduction

1.1 Perceptual Image Processing

Perception may roughly be defined as the mind's process of disentangling information from the physical means by which it is conveyed. As such, the study of perception is inherently a study of abstract representations. In distilling information into abstract quanta, the human mind prepares such information for conscious processing and consumption. In *visual* perception, the inputs are retinal images and the outputs are mental descriptions (e.g., shape and material properties) of the subject being observed.

On a spectrum of scales at which to probe human perception, the traditional endeavors of cognitive psychologists and visual system neuroscientists may coarsely be described as sitting at the two extremes. The former examines perceptual issues on a qualitative and macroscopic level while the latter identifies the microscopic biological building

blocks that allow for physical realization. Computational vision glues these two ends of the spectrum together by providing an algorithmic description of how the functional building blocks may give rise to the qualitative behaviors observed and classified through cognitive experiments. Such a process (i.e., the algorithmic transformation of information) can be studied independent of any of its physical instantiations, and hence falls within the purview of computer science.

The aim of this dissertation is to propose new algorithmic models for aspects of perceptual visual processing. While these models should replicate important features of human vision, we by no means expect our models to fully describe reality. Rather, they are simply meant to provide “first order” quantitative predictions; such predictions can then be used to either bolster theories or point out assumptions in need of greater refinement. *One* of our hopes is that these models might inform future experimental designs and aid (even if only as a distracting, but somewhat informative, detour) in science’s ongoing and successive approximations to truth.

Ultimately, however, our primary aim in this work is not to predict actual human behavior. Our main interest is in the design of new algorithms that assist computers in attacking various problems in graphics and vision. Motivation for this comes from the robust manner in which the humans cope with environmental variations and successfully complete a variegated set of visual tasks. By constructing new perception-inspired algorithms, we may be able to endow computers with comparable robustness in accomplishing similar goals. Therefore, the utility of these models will be judged by their usefulness in algorithmic problem solving; their value (as presented here) stands independent of the goodness of our assumptions on how humans actually process visual information. In this work, we

seek progress on three problems: achieving color constancy, designing an “illumination-invariant” color space, and relating projected image data to three-dimensional shape understanding.

1.1.1 Color Constancy

Color constancy refers to the fact that the perceived color of an object in a scene tends to remain constant even when the spectral properties of the illuminant (and thus the tristimulus data recorded at the retina) are changed drastically.

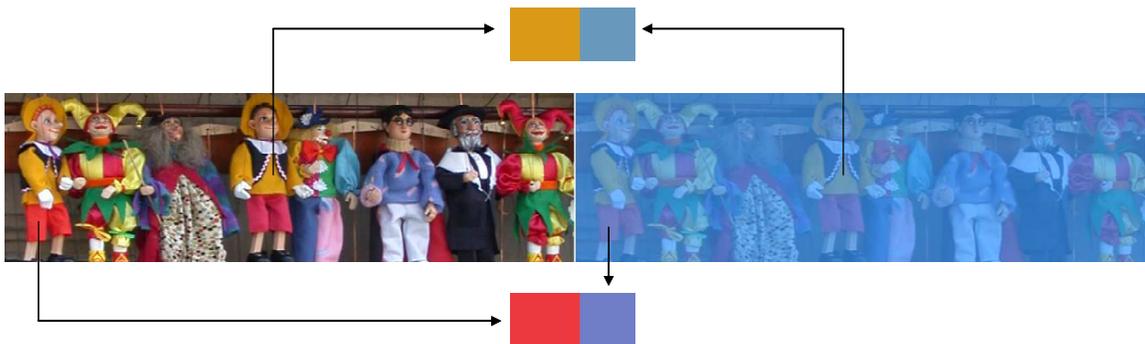


Figure 1.1: Arrows indicate pixel intensity values that have been replicated as uniform squares of color outside the image context for side-by-side comparison with pixel intensity values from the other image.

Figure 1.1 illustrates a local version of the effect. The left half of the figure shows a set of puppets lit by normal daylight. The right half of the figure shows the same set of puppets, but after an extreme environmental manipulation has been applied. While no human viewer would consider pairs of spatially corresponding pixels between left and right halves to be sharing the same color, even under this extreme change, the qualitative experience of the colors is mostly preserved. For example, if asked to coarsely label the

color of the center puppet's shirt on the right (marked by one of the arrows), one would settle upon the descriptor "yellow." However, when the pixel RGB values are pulled out and displayed against a white background, one instead labels the same RGB values as "blue." The surprising fact is that we do not simply see the right half of Figure 1.1 as consisting only of various shades of blue or purple. We still experience, for example, sensations of yellow (which is widely considered in vision science the color most opposite that of blue [46]).

Color constancy is even more powerful when the same environmental manipulation is applied to our entire visual field (i.e., is applied globally rather than locally). In such cases, we do not have the bright white of the paper (or the original image colors) to deduce the overall "blue-ness" of the new environment. For a common, simulated example of such a global manipulation, consider the eventual color shift that occurs after putting on a pair of ski goggles. The act of putting on goggles can be thought of as simulating a change of illuminant to one in which there is less power in the short wavelengths. At first, putting on a pair of ski goggles with an "orange" filter makes the world appear orange. However, over time, the orange-ness disappears and colors are largely perceived as if goggles had not been worn. When the goggles are later taken off, the world immediately appears exceedingly blue until color perception once again shifts into an adapted state.

This phenomenon results from some kind of global processing that is still not completely understood. One working model for achieving constancy is that the human first somehow estimates the illuminant. The effect of illumination is then undone by applying a "tone mapping" operator. This yields colors that are perceived as being lit under some canonical illuminant. This interpretation may be referred to as *illumination discounting*

[18]. Another working model is that the human processes retinal images by relating colors of observed objects. In converting retinal signals into perceived responses, the brain only uses relationships that are not affected by illumination changes. Such an interpretation may be referred to as *spatial processing* [18,43].

Whatever the mechanism, this perceptual processing allows color to be treated as an intrinsic material property. From our day-to-day perceptual experiences, we accept it as sensible to refer to an apple as red, or an orange as orange. This, however, belies the fact that color constancy is actually hard to achieve. Recall that the light seen reflected from objects depends on the full spectrum of the illuminant and the per-wavelength attenuation due to the material. Light is only turned into three sensory responses once it hits the eye. Therefore, given scenes with arbitrary illuminant spectra and material reflectances, there is no reason that constancy should be even close to possible. Consider the scenario in which two different materials appear the same under one illuminant (i.e., they are *metamers*), but look different under a second illuminant. In this case, one color maps to two distinct colors with illuminant change, so color relations clearly do not remain constant.

Despite the manifest difficulty in achieving color constancy, models of color constancy are often furthermore forced to obey an even more stringent constraint, the generalized von Kries hypothesis [21]. This hypothesis stipulates that color correction is done using independent scalings of the three tristimulus values. Von Kries' commonly adopted model assumes that the effect of changing illumination is to multiply each of our trichromatic sensor responses by a separate (possibly spatially varying) scale factor, and the brain is thus able to discount the change by applying the inverse scale factor [18, 33, 46]. The *generalized* von Kries model permits a change of color basis to take place before the gain

factors are applied.

With such demanding restrictions, color constancy may seem a vain endeavor. Indeed, humans do not have perfect color constancy; however, our color percepts are nonetheless surprisingly stable [46]. This suggests that the world's illuminants and reflectances are not completely arbitrary. One natural question is then, "What are the necessary and sufficient conditions under which constancy is in fact achievable?" Answers to this are discussed in Chapter 2. There, we will focus mostly on the generalized von Kries model. A related question is, "To what extent are these conditions true in our everyday world?" We will present experimental tests of this as well. Note that these questions concern the general *possibility* of achieving color constancy with *any* von Kries-like approach. Therefore, answers to these questions provide blanket statements that upper bound what is attainable for any method employing per-channel gain control for color constancy (e.g., grey world, grey edge [54], gamut mapping [24], retinex [33]). They do not address what algorithm in particular best achieves color constancy under these conditions.

Given that the necessary conditions for generalized von Kries models to achieve perfect color constancy may not be met exactly, we would also like to compute an optimal basis in which to run von Kries-based color constancy algorithms. The problem of computing such a basis is addressed in Section 2.3. As a follow-up question we also pose the following: "To what extent is there evidence that humans process under the basis we compute?" This question is left unanswered in this thesis.

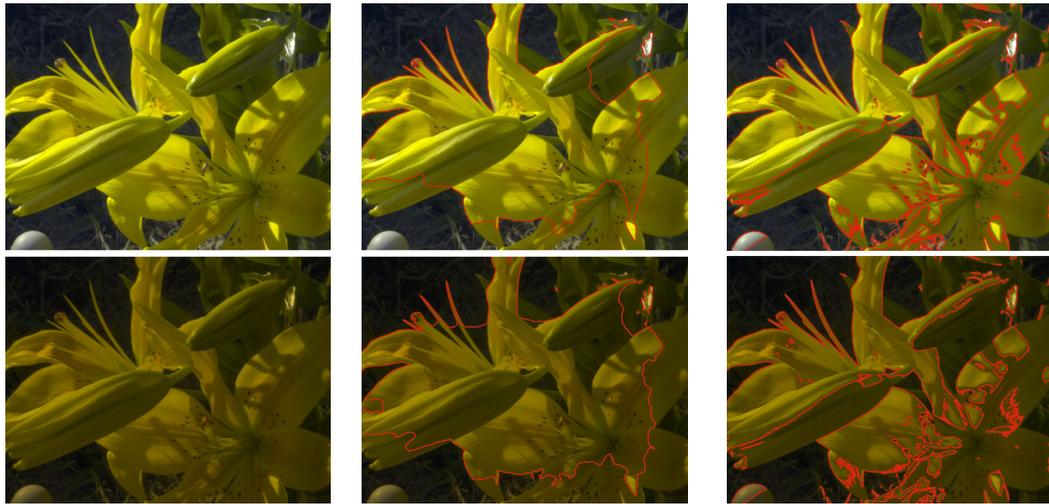
With regards to applications, models for color constancy can be applied to color correction problems such as white balancing. In white balancing, a scene is captured under

one ambient illuminant. The image is then later displayed, but under viewing conditions that employ a different ambient illuminant. The mismatch in illuminants during capture and display, and hence mismatch in observer adaptations in both cases, causes discrepancies in the perceived colors of the displayed reproduction. A white balanced image is one in which the captured colors are mapped to colors as seen under the new display illuminant. By displaying the white balanced image, a viewer perceives the reproduced scene in the same way the one who captured the image perceived the original scene.

1.1.2 A Perception-based Color Space

Given the importance of color processing in computer graphics, color spaces abound. While existing color spaces address a range of needs, they all suffer from one notable problem that makes them unsatisfactory for a large class of applications including segmentation and Poisson image editing [48]: algorithms working in these color spaces exhibit great sensitivity to the illumination conditions under which their inputs were captured.

Consider Figure 1.2 in which an active contour segmentation algorithm is run on two images that differ only in the scene's illumination (these were generated using full-spectral images and measured illuminant spectra). Figure 1.2a shows the two initial images. Figure 1.2b shows their segmentations in a typical color space (the space of RGB values provided by the image). Figure 1.2c shows their segmentations in another common color space (CIE $L^*a^*b^*$). In both color spaces, the segmentation parameters were tuned on each image to produce as clean and consistent segmentations as possible. Furthermore, even after parameter tweaking, the segmentations do not appear particularly reliable. It would



(a) Pair of images to segment. (b) One typical color space. (c) Another color space.

Figure 1.2: Active contour segmentation of two images in two different color spaces with parameters tweaked for each image so that segmentations are as clean and consistent as possible. Optimal parameters for the image pairs are separated by 2 and 5 orders of magnitude respectively.

be beneficial to have a segmentation algorithm that could use the same set of parameters to produce consistent segmentations of the two images in spite of their differing illuminations.

Figure 1.3 shows another example, this time of Poisson image editing. Here, a picture of a bear is cut out from one image (Figure 1.3a) and inserted into a background image of a swimming pool (Figure 1.3b). Figure 1.3c shows the result when run in a usual color space (the space of RGB values provided by the image). In this image, the foreground object is seamlessly inserted into the background image; however, the bear appears rather faded and ghostly. For comparison, Figure 1.3d shows an insertion performed using the color space we present in Chapter 3. Here the bear appears much more corporeal.

These examples point to the fact that algorithms working in current color spaces lack an essential robustness to the illumination conditions of their inputs. If processing is meant to operate on intrinsic scene properties, then such illumination-dependent results

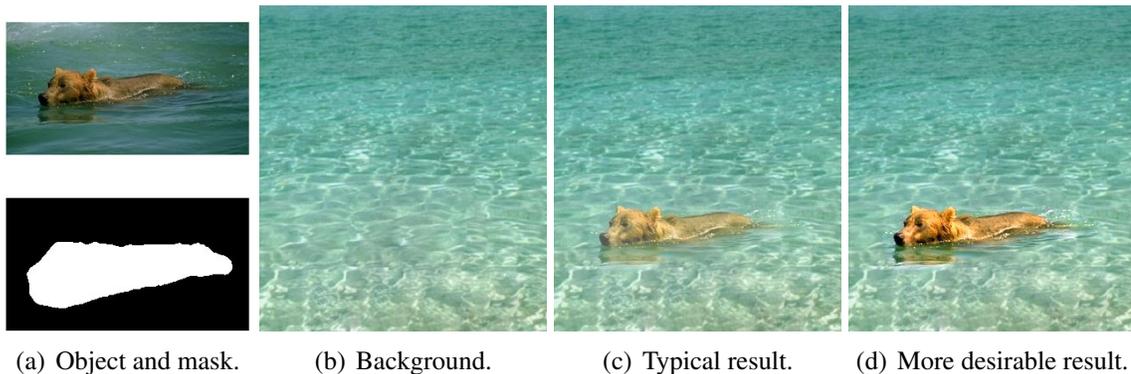


Figure 1.3: Poisson image editing.

are unsatisfactory. Since color constancy deals with taking measured colors and turning them into stable percepts, addressing this problem is plausibly related to issues of color constancy. Given an algorithm for color constancy, one solution might be to map an image's input colors to some canonical set of colors (i.e., colors as seen under a standard illuminant) and then work in such a space instead. One potential problem with such an approach is that it may depend on the ability of the algorithm to perform difficult tasks like estimating the illuminant in order to map colors to their canonical counterparts. An even better scenario would be if some ability to edit colors in an intrinsic sense were built into the choice of color space itself—without requiring the adoption of a specific color correction algorithm.

1.1.3 Shape Perception and Line Drawings

Shape perception refers to the process of taking image data and inferring the three-dimensional geometry of objects within the image. While there are many approaches to estimating shape (e.g., shape from shading [29], shape from texture [5], shape from specular flow [2]), we shall be interested in the connection between shape understanding and curves in an image.

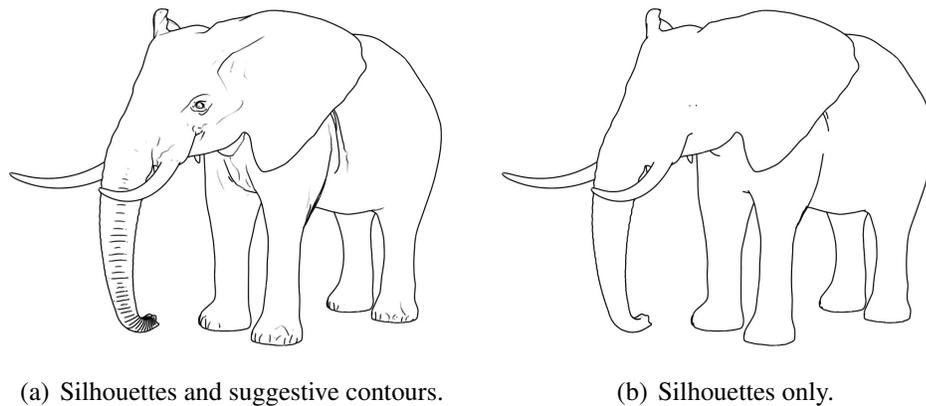


Figure 1.4: Line renderings of an elephant.

As shown in Figure 1.4a, line drawings alone are often enough to convey a sense of shape. This suggests that curves themselves provide a lot of information about a surface. Artists are able to make use of this fact and invert the process; they take a reference object, either observed or imagined, and summarize it as a line drawing that is almost unambiguously interpretable. Two natural questions then emerge: (1) Given a three-dimensional object, what curves should one draw to best convey the shape? (2) Given an image, is there a stable set of curves that humans can robustly detect and use to infer shape—and if so, how are they defined and what precisely is their information content?

Analyses of the more classically defined curves on surfaces (e.g., silhouettes, ridges, parabolic lines) are inadequate in giving a precise understanding to this problem. For example, Figure 1.4b shows that the relationships between curves and shape perception involve much more than simply silhouette data. Likewise can be said of the traditionally defined ridges and valleys. Researchers have more recently defined curves such as suggestive contours [15] and apparent ridges [32] which take advantage of view information and produce informative renderings, but more precise analyses of when and why they succeed

are still lacking.

A complete understanding of the set of curves humans latch onto clearly extends far beyond the scope of this work; we instead focus on laying some groundwork for future investigations. Our guiding philosophy is roughly as follows: humans only have retinal images as inputs; assuming curve-based shape understanding is a learned process, any meaningful curves should then be detectable (in a stable manner) in retinal images; therefore, analysis should be developed for relating curves defined in images to geometric properties of the viewed surface. Philosophy aside, the same geometric techniques developed for studying such relations may also be applied to answering other sorts of questions relating curves on surfaces to curves in images, so we hope these techniques may prove useful more generally.

1.2 Contributions

Color Constancy. Color constancy is almost exclusively modeled with von Kries (i.e., diagonal) transforms [21]. However, the choice of basis under which von Kries transforms are taken is traditionally *ad hoc*. Attempts to remedy the situation have been hindered by the fact that no joint characterization of the conditions for worlds¹ to support (generalized) von Kries color constancy has previously been achieved. In Chapter 2, we establish the following:

¹By a “world,” we mean a collection of illuminants, reflectances, and visual sensors. These are typically described by a database of the spectral distributions for the illuminants, per-wavelength reflections for the materials, and per-wavelength sensitivities for the sensors.

- Necessary and sufficient conditions for a world to support doubly linear color constancy.
- Necessary and sufficient conditions for a world to support generalized von Kries color constancy.
- An algorithm for computing a locally optimal color basis for generalized von Kries color constancy.

In the applications we discuss, we are mostly concerned with the cone sensors of the human visual system. However, the same theory applies for camera sensors and can be generalized in a straightforward manner for a different number of sensors.

A Perception-based Color Space. In Chapter 3, we design a new color space with the following properties:

- It has a simple 3D parameterization.
- Euclidean distances approximately match perceptual distances.
- Color displacements and gradients are approximately preserved under changes in the spectrum of the illuminant.
- Coordinate axes have an interpretation in terms of color opponent channels.
- It can easily be integrated with fibered space and measure-theoretic image models.

The first three bullet points imply the color space can easily be plugged into many image processing algorithms (i.e., algorithms can be called as black-box functions and require no change in implementation). This color space also enables these algorithms to work on the

intrinsic image and thereby yield (approximately) illumination-invariant results.

Shape Perception and Line Drawings. To study shape, we present some technical tools in the context of relating Saint-Venant curves to suggestive contours. Contributions include the following:

- Choice of bases for surface and image that simplifies calculations and makes expressions more conducive to interpretation.
- Transplanting of techniques for working with bases that change at every point and notation for reducing ambiguities.
- General procedure for expressing image information in terms of three-dimensional geometric quantities.
- Some relations between suggestive contours, Saint-Venant curves, and surface geometry.
- Relation between a surface-curve's normal curvature in the tangent direction and its apparent curvature in the image.

1.3 Overview of Thesis

The organization for the rest of the thesis is as follows:

Chapter 2. We analyze the conditions under which various models for color constancy can work. We focus in particular on the generalized von Kries model. We observe that the von

Kries compatibility conditions are impositions only on the *sensor measurements*, not the *physical spectra*. This allows us to formulate the conditions succinctly as rank constraints on an order three measurement tensor. Given this, we propose an algorithm that computes a (locally) optimal choice of color basis for von Kries color constancy and compare the results against other proposed choices.

Chapter 3. Motivated by perceptual principles—in particular, color constancy—we derive a new color space in which the associated metric approximates perceived distances and color displacements capture relationships that are robust to spectral changes in illumination. The resulting color space can be used with existing image processing algorithms with little or no change to the methods. We show application to segmentation and Poisson image processing.

Chapter 4. This chapter presents the mathematical background for Chapter 5 (where we consider the relations between shapes and image curves). The previous chapters make only minor references to this chapter, so readers only interested in the color processing portions of this work can simply refer to this chapter as required. Chapter 2 makes use of multilinear algebra, so reading the notation used for denoting tensors is sufficient. Chapter 3 makes reference to the metric equivalence problem discussed in Section 4.3. The more involved details are not so important, so a high level overview suffices. Chapter 5 makes full use of the presented differential geometry. So readers interested in that chapter should read Chapter 4 with more care.

Chapter 5. We use the formalism presented in Chapter 4 to relate curves detectable in images to information about the geometry. We focus in particular on Saint-Venant valleys

and relate them to suggestive contours. We develop techniques for relating image and surface information and apply these to characterizing curve behavior at critical points of illumination. We also prove more limited results away from critical points of illumination.

Appendix A. This appendix provides proofs for the results cited in Chapter 2. It also contains an analysis of the structure of the space of worlds supporting generalized von Kries color constancy.

Appendix B. This appendix presents the proof that locks down the functional form of our color space parameterization. We also prove that our model recovers the well-known Weber's Law for brightness perception.

Appendix C. This appendix contains more detailed discussions on differential geometry that are somewhat tangential to the main exposition. It includes some calculations that can be used to verify some of the claims or get a better sense for how the formalism can be used.

Appendix D. This appendix provides calculations and proofs for the various relations we discuss in Chapter 5. It also details some further derivations that may be useful for future work.

Chapter 2

A Basis for Color Constancy

In this chapter we investigate models for achieving color constancy. We devote particular attention to the ubiquitous generalized von Kries model. In Section 2.2.1, we relate the ability to attain perfect color constancy under such a model to the rank of a particular order three tensor. For cases in which perfect color constancy is not possible, this relationship suggests a strategy for computing an optimal color basis in which to run von Kries based algorithms.

2.1 Introduction and Previous Work

For a given scene, the human visual system, post adaptation, will settle on the same perceived color for an object despite spectral changes in illumination. Such an ability to discern illumination-invariant material descriptors has clear evolutionary advantages and also

largely simplifies (and hence is widely assumed in) a variety of computer vision algorithms.

To achieve color constancy, one must discount the effect of spectral changes in the illumination through transformations of an observer's trichromatic sensor response values. While many illumination-induced transformations are possible, it is commonly assumed that each of the three sensors reacts with a form of independent gain control (i.e., each sensor response value is simply scaled by a multiplicative factor), where the gain factors depend only on the illumination change [21, 33]. This is termed von Kries adaptation. Represented in linear algebra, it is equivalent to multiplying each column vector of sensor response values by a *shared* diagonal matrix (assuming spatially uniform illumination), and is therefore also referred to as the diagonal model for color constancy.

Note that while the initial von Kries hypothesis applied only to direct multiplicative adjustments of retinal cone sensors, we follow [21] and use the term more loosely to allow for general trichromatic sensors. We also allow for a change of color basis to occur before the per-channel multiplicative adjustment. (Finlayson et al. [21] refer to this as a *generalized diagonal* model for color constancy, and they term the change of color basis a *sharpening transform*.)

The (generalized) diagonal model is at the core of the majority of color constancy algorithms. Even a number of algorithms not obviously reliant on the diagonal assumption in fact rely on diagonal models following a change of color basis [21, 22]; their choice of color basis is simply not explicit. Yet, despite the widespread use of the diagonal model, good choices of color bases under which diagonal transforms can be taken are only partially understood.

The most theoretically-justified approach to choosing a color space is predicated on the assumption that the spaces of illuminant and reflectance spectra are each low-dimensional. As shown by Finlayson et al. [21, 22], a two-dimensional linear space of illuminants and a three-dimensional linear space of reflectances (or vice versa) is sufficient¹ to guarantee diagonal color constancy. This is an important observation because it provides a principled method for choosing a von-Kries compatible color space for a given set of sensors, illuminants, and reflectances.

While the work of Finlayson et al. [21,22] is a significant first step, both empirical and analytical observations reveal this characterization to be incomplete. Empirical studies suggest that diagonal color constancy is adequate under more general conditions than the 2–3 model implies [20]. (This is good news for vision systems that seek color constancy in the real world. While the space of natural illuminants may be sufficiently small, the space of reflectances is probably not [41].)

From an analytical standpoint, the 2–3 model is known to be only sufficient¹—not necessary—for von Kries adaptation to work. Indeed, once a specific color basis is chosen, one can expand the set of compatible illuminants well beyond a two-dimensional space by adding those illuminants that integrate to zero against (i.e., are black with respect to) all reflectances in the compatible set [20]. More surprisingly, one can algorithmically grow the compatible illuminant set to include additional *non-black* illuminants as well [20, 58]. This analysis, however, is still incomplete because given color data, the analysis does not reveal a method for computing the color space to begin with.

¹For sufficiency, we must allow complex color bases. See section 2.4

While these limitations have been well-documented, a more complete characterization of the conditions for von Kries compatibility has yet to be established. As a result, the development of more powerful systems for choosing optimized color bases has been slow. This chapter addresses these issues by answering the following questions:

- (1) What are the necessary and sufficient conditions that sensors, illuminants, and materials must satisfy to be exactly von Kries compatible, and what is the structure of the solution space?
- (2) Given measured spectra or labeled color observations, how do we determine the color space that “best” supports diagonal color constancy?

We observe that the joint conditions are impositions only on the *sensor measurements*, not the *physical spectra*. This allows the von Kries compatibility conditions to be succinctly formulated as rank constraints on an order 3 measurement tensor. Our analysis leads directly to an algorithm that, given labeled color data, computes a locally optimal choice of color basis in which to carry out diagonal color constancy computations. The proposed framework also unifies most existing analyses of von Kries compatibility.

2.2 Theory

We define two notions of color constancy. The first definition captures the idea that a single adjustment to the sensors will map all material colors seen under an illuminant E_1 to reference colors under (a possibly chosen standard) illuminant E_2 . The second definition (also known as *relational* color constancy) captures the idea that surface colors have a fixed

relationship between each other no matter what overall illumination lights the scene. As stated, these two definitions are not interchangeable. One being true does not imply the other.

To define the issues formally, we need a bit of notation. Let \mathcal{R} be the smallest (closed) linear subspace of L_2 functions enclosing the spectral space of materials of interest. Let \mathcal{E} be the smallest (closed) linear subspace of L_2 functions enclosing the spectral space of illuminants of interest. Let $\vec{p}^{R,E}$ be the color (in the sensor basis) of material reflectance $R(\lambda) \in \mathcal{R}$ under illumination $E(\lambda) \in \mathcal{E}$. In the following, D and \bar{D} are operators that take color vectors and map them to color vectors. D is required to be independent of the material R ; likewise, \bar{D} is required to be independent of the illuminant E . The $*$ denotes the action of these operators on color vectors.

(1) Color constancy:

For all $E_1, E_2 \in \mathcal{E}$, there exists a $D(E_1, E_2)$ such that for all $R \in \mathcal{R}$,

$$\vec{p}^{R,E_2} = D(E_1, E_2) * \vec{p}^{R,E_1}$$

(2) Relational color constancy:

For all $R_1, R_2 \in \mathcal{R}$, there exists a $\bar{D}(R_1, R_2)$ such that for all $E \in \mathcal{E}$,

$$\vec{p}^{R_2,E} = \bar{D}(R_1, R_2) * \vec{p}^{R_1,E}$$

In the case that D and \bar{D} are linear (and hence identified with matrices), $*$ is just matrix-vector multiplication. If D is linear, we say that the world supports *linear adaptive* color constancy. If \bar{D} is linear, we say the world supports *linear relational* color constancy. D being linear does not imply \bar{D} is linear, and vice versa. If both D and \bar{D} are linear, we say the world supports *doubly linear* color constancy.

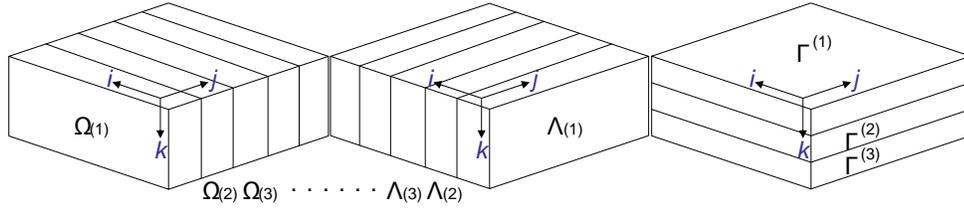


Figure 2.1: The $3 \times I \times J$ measurement tensor. The tensor can be sliced in three ways to produce the matrices $\Omega_{(j)}$, $\Lambda_{(i)}$, and $\Gamma^{(k)}$.

In particular, we shall be interested in the case when D and \bar{D} are both furthermore *diagonal* (under some choice of color basis). It is proven in [21] that for a fixed color space, D is diagonal if and only if \bar{D} is diagonal. So the two notions of color constancy are equivalent if either D or \bar{D} is diagonal, and we say the world supports *diagonal* color constancy (the *doubly* modifier is unnecessary). The equivalence is nice because we, as biological organisms, can likely learn to achieve definition 1, but seek to achieve definition 2 for inference.

Given a set of illuminants $\{E_i\}_{i=1,\dots,I}$, reflectances $\{R_j\}_{j=1,\dots,J}$, and sensor color matching functions $\{\rho^k\}_{k=1,2,3}$, we define a measurement data tensor² (see Figure 2.1):

$$M^k_{ij} := \int \rho^k(\lambda) E_i(\lambda) R_j(\lambda) d\lambda \quad (2.1)$$

For fixed values of j , we get $3 \times I$ matrices $\Omega_{(j)} := M^k_{ij}$ that map illuminants expressed in the $\{E_i\}_{i=1,\dots,I}$ basis to color vectors expressed in the sensor basis. Likewise, for fixed values of i , we get $3 \times J$ matrices $\Lambda_{(i)} := M^k_{ij}$ that map surface reflectance spectra expressed in the $\{R_j\}_{j=1,\dots,J}$ basis to color vectors. We can also slice the tensor by constant

²We will use latin indices starting alphabetically with i to denote tensor components instead of the usual greek letters to simplify notation and avoid confusion with other greek letters floating around.

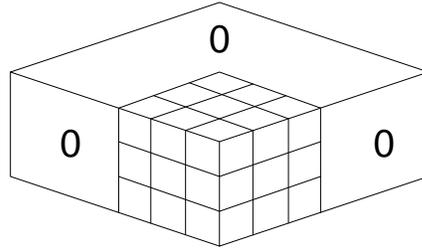


Figure 2.2: Core tensor form: a $3 \times 3 \times 3$ core tensor is padded with zeros. The core tensor is not unique.

k to get $I \times J$ matrices $\Gamma^{(k)} := M^k_{ij}$.

Since color perception can depend only on the eye's trichromatic color measurements, worlds (i.e., sets of illuminant and material spectra) giving rise to the same measurement tensor are perceptually equivalent. To understand diagonal color constancy, therefore, it is sufficient to analyze the space of measurement tensors and the constraints that these tensors must satisfy. This analysis of von Kries compatible measurement tensors is covered in section 2.2.1.

Given a von Kries compatible measurement tensor (e.g., an output from the algorithm in section 2.3), one may also be interested in the constraints such a tensor places on the possible spectral worlds. This analysis is covered in section 2.4.

2.2.1 Measurement Constraints

The discussion in this section will always assume generic configurations (e.g., color measurements span three dimensions, color bases are invertible). Proofs not essential to the main exposition are relegated to Appendix A.

Proposition 1. *A measurement tensor supports doubly linear color constancy if and only*

if there exists a change of basis for illuminants and materials that reduces it to the core tensor form of Figure 2.2.

More specifically (as is apparent from the proof of Proposition 1 in Appendix A.1.1), if a single change of illuminant basis makes all the $\Omega_{(j)}$ slices null past the third column, the measurement tensor supports linear relational color constancy. Likewise, a change of material basis making all the $\Lambda_{(i)}$ slices null past the third column implies the measurement tensor supports linear adaptive color constancy. Support for one form of linear constancy does not imply support for the other.

The following lemma provides a stepping stone to our main theoretical result and is related to some existing von Kries compatibility results (see section 2.4).

Lemma 1. *A measurement tensor supports generalized diagonal color constancy if and only if there exists a change of color basis such that, for all k , $\Gamma^{(k)}$ is a rank-1 matrix.*

This leads to our main theorem characterizing the space of measurement tensors supporting generalized diagonal color constancy.

Theorem 1. *A measurement tensor supports generalized diagonal color constancy if and only if it is a rank 3 tensor.³*

An order 3 tensor (3D data block) T is rank N if N is the smallest integer such that there exist vectors⁴ $\{\vec{a}_n, \vec{b}_n, \vec{c}_n\}_{n=1, \dots, N}$ allowing decomposition as the sum of outer products (denoted by \circ):

³There exist measurement tensors supporting generalized diagonal color constancy with rank less than 3, but such examples are not generic.

⁴In the language of differential geometry introduced in Chapter 4, $\{\vec{a}_n, \vec{b}_n\}_{n=1, \dots, N}$ are really covector component lists. $\{\vec{c}_n\}_{n=1, \dots, N}$ are really vector component lists.

$$T = \sum_{n=1}^N \vec{c}_n \circ \vec{a}_n \circ \vec{b}_n. \quad (2.2)$$

Without loss of generality, let $\{\vec{a}_n\}$ be vectors of length I , corresponding to the illuminant axis of the measurement tensor; let $\{\vec{b}_n\}$ be vectors of length J , corresponding to the material axis of the tensor; and let $\{\vec{c}_n\}$ be vectors of length 3, corresponding to the color sensor axis of the tensor. Let the vectors $\{\vec{a}_n\}$ make up the columns of the matrix A , vectors $\{\vec{b}_n\}$ make up the columns of the matrix B , and vectors $\{\vec{c}_n\}$ make up the columns of the matrix C . Then the decomposition above may be restated as a decomposition into the matrices (A, B, C) , each with N columns.

Proof. (Theorem 1). First suppose the measurement tensor supports generalized diagonal color constancy. Then by Lemma 1, there exists a color basis under which each $\Gamma^{(k)}$ is rank-1 (as a matrix). This means each $\Gamma^{(k)}$ can be written as an outer product, $\Gamma^{(k)} = \vec{a}_k \circ \vec{b}_k$. In this color basis then, the measurement tensor is a rank 3 tensor in which the matrix C (following notation above) is just the identity.⁵ We also point out that an invertible change of basis (on any of A, B, C) does not affect the rank of a tensor, so the original tensor (before the initial color basis change) was also rank 3.

For the converse case, we now suppose the measurement tensor is rank 3. Since C is (in the generic setting) invertible, multi-linearity gives us

$$C^{-1} * \left(\sum_{n=1}^3 \vec{c}_n \circ \vec{a}_n \circ \vec{b}_n \right) = \sum_{n=1}^3 (C^{-1} \vec{c}_n) \circ \vec{a}_n \circ \vec{b}_n. \quad (2.3)$$

⁵Technically, this shows that the tensor is at most rank 3, but since we are working with generic tensors, we can safely discard the degenerate cases in which observed colors do not span the three-dimensional color space.

The operator on the left hand side of Equation (2.3) denotes the application of the 3×3 matrix C^{-1} along the sensor axis of the tensor. The right hand side of Equation (2.3) is a rank 3 tensor with each $\Gamma^{(k)}$ slice a rank-1 matrix. By Lemma 1, the tensor must then support diagonal color constancy. ■

In the proof above, note that the columns of C exactly represent the desired color basis under which we get perfect diagonal color constancy. This theorem is of algorithmic importance because it ties the von Kries compatibility criteria to quantities (best rank 3 tensor approximations) that are computable via existing multilinear methods.

2.3 Color Basis for Color Constancy

Given a measurement tensor M generated from real-world data, we would like to find the optimal basis in which to perform diagonal color constancy computations. To do this, we first find the closest von Kries compatible measurement tensor (with respect to the Frobenius norm). We then return the color basis that yields perfect color constancy under this approximate tensor.

By Theorem 1, finding the closest von Kries compatible measurement tensor is equivalent to finding the best rank 3 approximation. Any rank 3 tensor may be written in the form of equation (2.2) with $N = 3$. It also turns out that such a decomposition of a rank three tensor into these outer-product vectors is almost always unique (modulo permutations and scalings). We solve for M 's best rank 3 approximation (decomposition into A, B, C) via Trilinear Alternating Least Squares (TALS) [28]. For a rank 3 tensor, TALS forces $A,$

B , and C to each have 3 columns. It then iteratively fixes two of the matrices and solves for the third in a least squares sense.

Repeating these computations in lockstep guarantees convergence to a local minimum. A , B , C can be used to reconstruct the closest von Kries compatible tensor and the columns of C exactly represent the desired color basis.

As a side note, the output of this procedure differs from the best rank-(3,3,3) approximation given by HOSVD [37]. HOSVD only gives orthogonal bases as output and the rank-(3,3,3) truncation does not in general yield a closest rank 3 tensor. HOSVD may, however, provide a good initial guess.

The following details on TALS mimic the discussion in [52]. For further information, see [28, 52] and the references therein. The Khatri-Rao product of two matrices A and B with N columns each is given by

$$A \odot B := \left[\vec{a}_1 \otimes \vec{b}_1, \vec{a}_2 \otimes \vec{b}_2, \dots, \vec{a}_N \otimes \vec{b}_N \right], \quad (2.4)$$

where \otimes is the Kronecker product.

Denote the flattening of the measurement tensor M by $\tilde{M}^{IJ \times 3}$ if the elements of M are unrolled such that the rows of matrix $\tilde{M}^{IJ \times 3}$ loop over the (i, j) -indices with $i = 1, \dots, I$ as the outer loop and $j = 1, \dots, J$ as the inner loop. The column index of $\tilde{M}^{IJ \times 3}$ corresponds with the dimension of the measurement tensor that is not unrolled (in this case $k = 1, 2, 3$).

The notation for other flattenings is defined symmetrically. We can then write

$$\tilde{M}^{JI \times 3} = (B \odot A) C^T. \quad (2.5)$$

By symmetry of equation (2.5), we can write out the least squares solutions for each of the matrices (with the other two fixed):

$$A = \left[(B \odot C)^\dagger \tilde{M}^{J3 \times I} \right]^T, \quad (2.6)$$

$$B = \left[(C \odot A)^\dagger \tilde{M}^{3I \times J} \right]^T, \quad (2.7)$$

$$C = \left[(B \odot A)^\dagger \tilde{M}^{JI \times 3} \right]^T. \quad (2.8)$$

2.4 Relationship to Previous Characterizations

As mentioned in the introduction, there are two main sets of theoretical results. There are the works of [20, 58] that give necessary and sufficient conditions for von Kries compatibility under a predetermined choice of color space, and are able to build infinite dimensional von Kries compatible worlds for this choice. Then there are the works of [21, 22] that prescribe a method for choosing the color space, but only for worlds with low dimensional linear spaces of illuminants and materials. We omit direct comparison to the various spectral sharpening techniques [6, 16, 23] in this section, as these methods propose more intuitive guidelines rather than formal relationships.

Previous analyses treat the von Kries compatibility conditions as constraints on

spectra, whereas the analysis here treats them as constraints on color measurements. In this section, we translate between the two perspectives. To go from spectra to measurement tensors is straightforward. To go the other way is a bit more tricky. In particular, given a measurement tensor with rank-1 $\Gamma^{(k)}$, there is not a unique world generating this data. Any set of illuminants $\{E_i\}_{i=1,\dots,I}$ and reflectances $\{R_j\}_{j=1,\dots,J}$ satisfying Equation (2.1) (with M and ρ_k fixed) will be consistent with the data. Many constructions of worlds are thus possible. But if one first selects particular illuminant or material spectra as mandatory inclusions in the world, then one can state more specific conditions on the remaining spectral choices. For more on this, see Appendix A.2.

In [21, 22], it is shown that if the illuminant space is 3 dimensional and the material space is 2 dimensional (or vice versa), then the resulting world is (generalized) von Kries compatible. As a measurement tensor, this translates into stating that any $3 \times 3 \times 2$ measurement tensor is (complex) rank 3. However this “3-2” condition is clearly not necessary as almost every rank 3 tensor is not reducible via change of bases to size $3 \times 3 \times 2$. In fact, one can always extend a $3 \times 3 \times 2$ tensor to a $3 \times 3 \times 3$ core tensor such that the $\Gamma^{(k)}$ are still rank-1. The illuminant added by this extension is neither black with respect to the materials, nor in the linear span of the first two illuminants.

The necessary and sufficient conditions provided in [58] can be seen as special cases of Lemma 1. The focus on spectra leads to a case-by-case analysis with arbitrary spectral preferences. However, the essential property these conditions point to is that the 2×2 minors of $\Gamma^{(k)}$ must be zero (i.e., $\Gamma^{(k)}$ must be rank-1).

One case from [58] is explained in detail in [20]. They fix a color space, a space

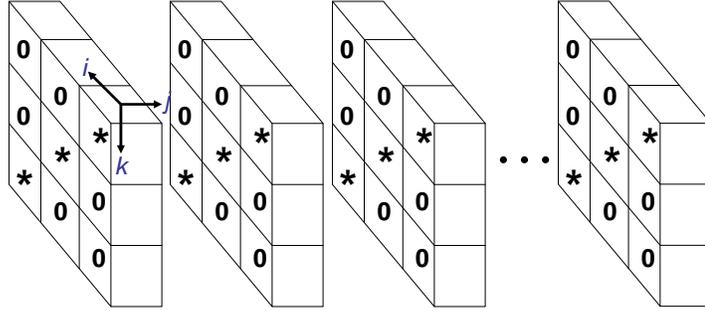


Figure 2.3: The rows of a single $\Lambda_{(1)}$ slice are placed into a new measurement tensor (rows are laid horizontally above) with all other entries set to zero. The * marks the nonzero entries.

of material spectra, and a single reference illumination spectrum. They can then solve for the unique space of illumination spectra that includes the reference illuminant and is von Kries compatible (in the fixed color basis) with the given material space.

In our framework, this can be interpreted as follows. The given input gives rise to a single $\Lambda_{(1)}$ measurement slice. The three rows of this slice can be pulled out and placed in a new measurement tensor of the form shown in Figure 2.3. This measurement tensor is then padded with an infinite number of zero $\Lambda_{(i)}$ matrices. The $\Gamma^{(k)}$ slices of this new tensor are clearly rank-1 matrices, and thus this tensor is von Kries compatible in the given color space. Moreover, any measurement tensor with rank-1 $\Gamma^{(k)}$ that include the original $\Lambda_{(1)}$ slice in its span must have $\Lambda_{(i)}$ slices that are spanned by the $\Lambda_{(i)}$ slices in Figure 2.3. With this fixed tensor and the fixed material spectra, one can then solve Equation (2.1) to obtain the space of compatible illumination spectra. This space can be described by three non-black illuminants and an infinite number of black illuminants (giving zero measurements for the input material space). Since the original $\Lambda_{(1)}$ measurement slice is in the span of the $\Lambda_{(i)}$ slices, the original reference illuminant must be in the solution space.

2.5 Results

We used the SFU color constancy dataset to create a measurement tensor to use in our experiments. The SFU database provides 8 illuminants simulating daylight, and 1,995 materials including measured spectra of natural objects. Fluorescent spectra were removed from the dataset in hopes of better modeling natural lighting conditions since, in color matching experiments, fluorescent lamps cause “unacceptable mismatches of colored materials that are supposed to match under daylight” [60].

Color matching functions were taken to be CIE 1931 2-deg XYZ with Judd 1951 and Vos 1978 modifications [60]. To resolve mismatches in spectral sampling, we interpolated data using linear reconstruction. Cone fundamentals were taken to be the Vos and Walraven (1971) fundamentals [60]. Experiments were run with illuminant spectra normalized with respect to the L_2 norm.

We followed the strategy of Section 2.3 to produce optimized color spaces for von Kries color constancy.

2.5.1 Effective Rank of “the World”

To test the “effective rank” of our world (as measured by available datasets), we approximate the SFU measurement tensor with tensors of varying rank and see where the dropoff in approximation error occurs. Figure 2.4 shows the results from the experiment. We measured error in CIEDE2000 units in an effort to match human perceptual error. CIEDE2000 is the latest CIE standard for computing perceptual distance and gives the best fit of any

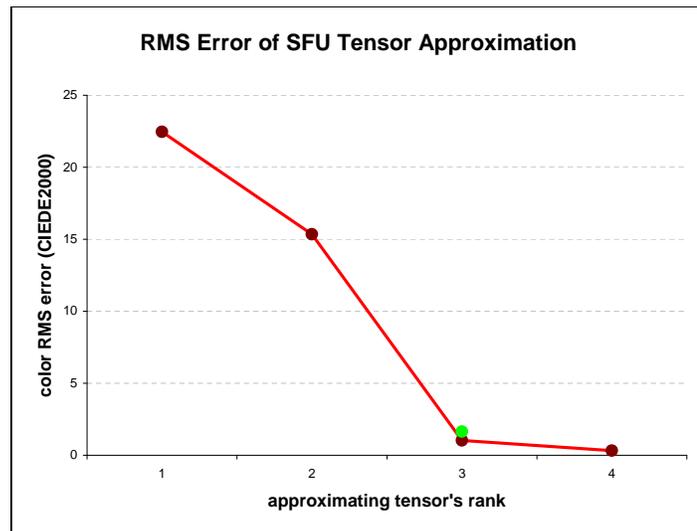


Figure 2.4: The effective rank of the SFU dataset is about 3. The green dot marks the error for the rank 3 tensor in which the color gamut was constrained to include the entire set of human-visible colors (see Section 2.5.5).

method to the perceptual distance datasets used by the CIE [40]. As a rule of thumb, 1 CIEDE2000 unit corresponds to about 1 or 2 just noticeable differences.

The red curve in Figure 2.4 shows that the SFU measurement tensor is already quite well approximated by a rank 3 tensor. The rank 2 tensor's approximation error is 68.3% of the rank 1 tensor's approximation error. The rank 3 tensor's approximation error is 6.7% of the rank 2 tensor's approximation error. The rank 4 tensor's approximation error is 31.3% of the rank 3 tensor's approximation error. We discuss the meaning of the green dot later in Section 2.5.5.

2.5.2 Von Kries Sensors

The matrix mapping XYZ coordinates to the new color coordinates, as computed using the TALS optimization on the SFU database, is given by the following (where the rows have

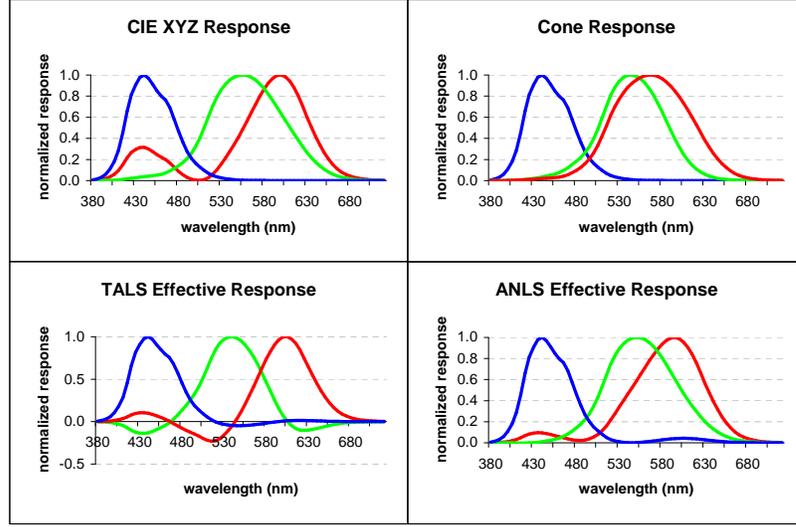


Figure 2.5: Color matching functions. Top row shows the standard CIE XYZ and Cone matching functions. The bottom row shows the effective sensor matching functions resulting from applying the C matrix derived from optimizations on the SFU database. ANLS (alternating nonlinear least squares) constrains the color space gamut to contain the human-visible gamut and measures perceptual error using CIEDE2000.

been normalized to unit length):

$$C^{-1} = \begin{bmatrix} 9.375111 \times 10^{-1} & -3.197979 \times 10^{-1} & -1.371214 \times 10^{-1} \\ -4.783729 \times 10^{-1} & 8.779015 \times 10^{-1} & 2.117135 \times 10^{-2} \\ 8.338662 \times 10^{-2} & -1.235156 \times 10^{-1} & 9.888329 \times 10^{-1} \end{bmatrix}. \quad (2.9)$$

We ran our algorithm on the Joensuu database as well [45]. The Joensuu database provides 22 daylight spectra and 219 natural material spectra (mostly flowers and leaves). The resulting basis vectors (columns of C) were within a couple degrees (in XYZ space) of the SFU optimized basis vectors. This seems to suggest some amount of stability in the result.

The effective sensor matching functions given by the optimized color basis are shown in Figure 2.5. As is intuitively suspected, the optimized basis causes a “sharpening”

in the peaks of the matching functions. This expectation is motivated by the fact that the diagonal model is exact for disjoint sensor responses. The “ANLS” result will be discussed in Section 2.5.5.

2.5.3 White Patch Normalization

There are many experiments one might devise to measure the color constancy afforded by various color bases. We chose to replicate a procedure commonly used in the literature. This experiment is based on a ‘white patch’ normalization algorithm, and is described below.

Dataset and algorithms

We ran our color basis algorithm on the SFU dataset [7] and compared our resulting color basis against previous choices (the cone sensor basis, 4 bases derived from different low dimensional approximations of spectral data, that of Barnard et. al. [6], and the “sensor sharpened” basis [23]).

The low dimensional worlds to which we compare are taken to have either 3 dimensional illuminant spaces and 2 dimensional material spaces (a 3-2 world) or vice versa (a 2-3 world); this allows computing color bases via the procedure in [21].

We took two different approaches to approximating spectra with low dimensional vector spaces. In the first approach (described in [21, 23]), we run SVD on the illuminant and material spectra separately. We then save the best rank-3 and rank-2 approximations.

This is Finlayson’s “perfect sharpening” method for databases with multiple lights [23], and is one of the algorithms that falls under the label of “spectral sharpening”.

As pointed out in [42], if error is to be measured in sensor space, there are alternatives to running PCA on spectra. Given a measurement tensor, the alternative (“tensor-based”) approach instead applies SVD on the tensor flattenings $\tilde{M}^{J3 \times I}$ and $\tilde{M}^{3I \times J}$ to get the principal combination coefficients of the spectral bases (to be solved for) that approximate the sample spectra. Refer to [42] for details.

We label the algorithm of Barnard et. al. [6] as “Barnard”. This is a more recent algorithm and also falls under the label of “spectral sharpening”.

We also ran experiments against Finlayson’s “sensor sharpening” algorithm. Note that this algorithm does not actually use the database in determining the sensor transforms. Its heuristic is simply to transform the sensors so that the responses are as “sharp” as possible.

We also tested against a modified version of Finlayson’s “database sharpening” method [23]. This algorithm as stated is defined in the case when the database has two illuminants and possibly many materials. Since our database has multiple lights, we used PCA on the set of illuminant spectra and ran the algorithm using the two most dominant principal components as our two illuminants. The results were nearly identical to both of the 2-3 methods, and we therefore omit the corresponding curves from the graphs.

Experimental procedure

We run the same white-patch normalization experiment as in [21]. As input, we are given a chosen white material W and an illuminant E . For the SFU database, we used the only material labeled as ‘white’ as our white material W . For every other material R , we compute a descriptor by dividing each of its 3 observed color coordinates by the 3 color coordinates of W (the resulting 3 ratios are then transformed as a color vector to XYZ coordinates so that consistent comparisons can be made with different choices of color space). In a von Kries world, the descriptor for R would not depend on the illuminant E . To measure the non von Kries-ness of a world, we can look at how much these descriptors vary with the choice of E .

More formally, we define the descriptor as:

$$\vec{d}_E^{W,R} = C [\text{diag} (C^{-1} \vec{p}^{W,E})]^{-1} C^{-1} \vec{p}^{R,E} \quad (2.10)$$

The function *diag* creates a matrix whose diagonal elements are the given vector’s components. C is a color basis. $C, \vec{p}^{R,E}, \vec{p}^{W,E}$ are given in the CIE XYZ coordinate system.

This means we compute the color vectors $\vec{p}^{W,E}$ and $\vec{p}^{R,E}$ as:

$$(\vec{p}^{W,E})^k := \int \rho^k(\lambda) E(\lambda) W(\lambda) d\lambda \quad (2.11)$$

$$(\vec{p}^{R,E})^k := \int \rho^k(\lambda) E(\lambda) R(\lambda) d\lambda \quad (2.12)$$

with $\rho^1(\lambda) = \bar{x}(\lambda) =$ CIE response function for X , $\rho^2(\lambda) = \bar{y}(\lambda) =$ CIE response function

for Y , and $\rho^3(\lambda) = \bar{z}(\lambda) =$ CIE response function for Z . The columns of the matrix C are the (normalized) basis vectors of a new color space in XYZ coordinates. C^{-1} is the inverse of C . This means that given some color vector \vec{v} in XYZ coordinates, $C^{-1}\vec{v}$ computes the coordinates of the same color vector expressed in the new color space's coordinate system.

To compute a non von Kries-ness error, we fix a canonical illuminant E' and compute descriptors $\vec{d}_{E'}^{W,R}$ for every test material R . We then choose some different illuminant E and again compute descriptors $\vec{d}_E^{W,R}$ for every test material R . Errors for every choice of E and R are computed as:

$$Error = 100 \times \frac{\|\vec{d}_{E'}^{W,R} - \vec{d}_E^{W,R}\|}{\|\vec{d}_{E'}^{W,R}\|} \quad (2.13)$$

For each instance of the experiment, we choose one SFU test illuminant E and compute the errors over all test materials (canonical illuminant E' is kept the same for each experimental instance). Each time, the color basis derived from running our method (labeled as “Optimized”) performed the best.

To visualize the results of the experimental test phase, we plot histograms of % color vectors that are mapped correctly with a diagonal transform versus the % allowable error. Hence, each histogram plot requires the specification of two illuminants—one canonical and one test. Figure 2.6 shows the cumulative histograms for instances in which the stated basis performs the best and worst relative to the next best basis. Relative performance between two bases is measured as a ratio of the areas under their respective histogram curves. The entire process is then repeated for another canonical E' to give a total of 4 graphs.

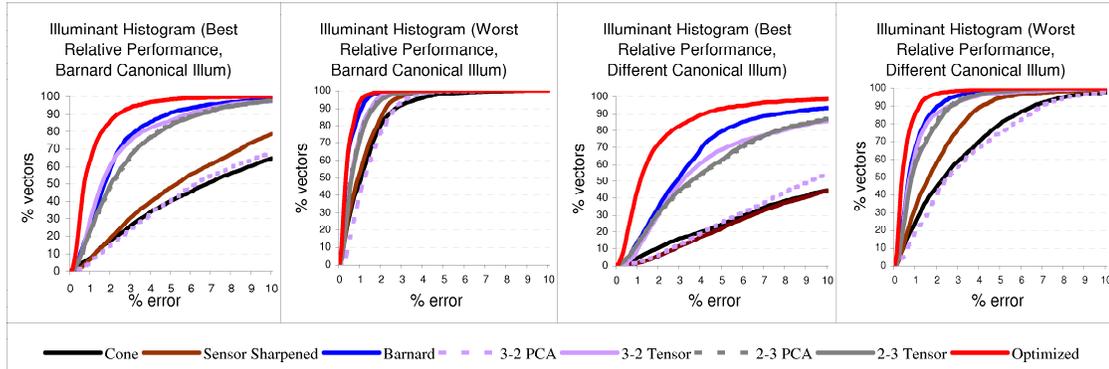


Figure 2.6: Percent vectors satisfying von Kries mapping versus percent allowable error. Each curve represents a different choice of color space. For low dimensional worlds, the dimension of the illuminant space precedes the dimension of the material space in the abbreviated notation. Low dimensional approximations were obtained either by running PCA on spectra or by tensor methods described in text. We show the experimental instances in which our derived basis performs the best and worst relative to the next best basis. The left and right halves differ in choice of canonical illuminant for testing. Unlike “Barnard”, our method effectively optimizes all pairs of lights.

White patch normalization results

The optimization algorithm labeled as “Barnard” requires specification of a canonical illuminant and chooses a color basis such that the mapping between any other test illuminant and the canonical one is as diagonal as possible (in a least squares sense).

To be fair, we show two sets of results: one where the canonical illuminant during test matches the canonical illuminant used in Barnard optimization; and one set where a different canonical illuminant is chosen during test from that used in Barnard optimization. The second canonical illuminant is chosen to illustrate the best worst-case relative performance of our algorithm. Goodness is measured as the ratio of areas under the histogram curves.

Barnard’s algorithm performs close to ours for some pairings of test and canoni-

cal illuminants, but is outperformed in most cases. These results are explained theoretically by the fact that, even though it optimizes over multiple light-pairs, the method of Barnard et al. always optimizes with respect to a single canonical illuminant. In contrast, we effectively optimize over all possible pairs of lights.

As noted earlier, we also tested against Finlayson’s “database sharpening” method [23] (using PCA on lights to handle multiple lights). The results were nearly identical to both of the 2-3 methods.

2.5.4 White Balancing

In this section, we discuss a rough perceptual validation we performed. We first downloaded a hyperspectral image from the database provided by [25]. Each pixel of the hyperspectral image gives a discretized set of samples for the per-wavelength reflectance function of the material seen at that pixel. The particular image we used was that of a flower shown in Figure 2.7. Assuming a lighting model in which viewed rays are simply computed as the product of the illuminant and reflectance spectra in the wavelength domain, we rendered the scene as it would appear under a tungsten light bulb and as it would appear under daylight. For visualization we converted all images into RGB coordinates and gamma corrected each coordinate before display. The goal of the algorithm was to start with the input image of the flower under a tungsten light and transform it into the target image of the flower under daylight using a generalized diagonal model. The goodness of a color basis is judged by the degree to which it facilitates the transformation of the input into the desired output (with the significance of discrepancies determined by human perceptual error).

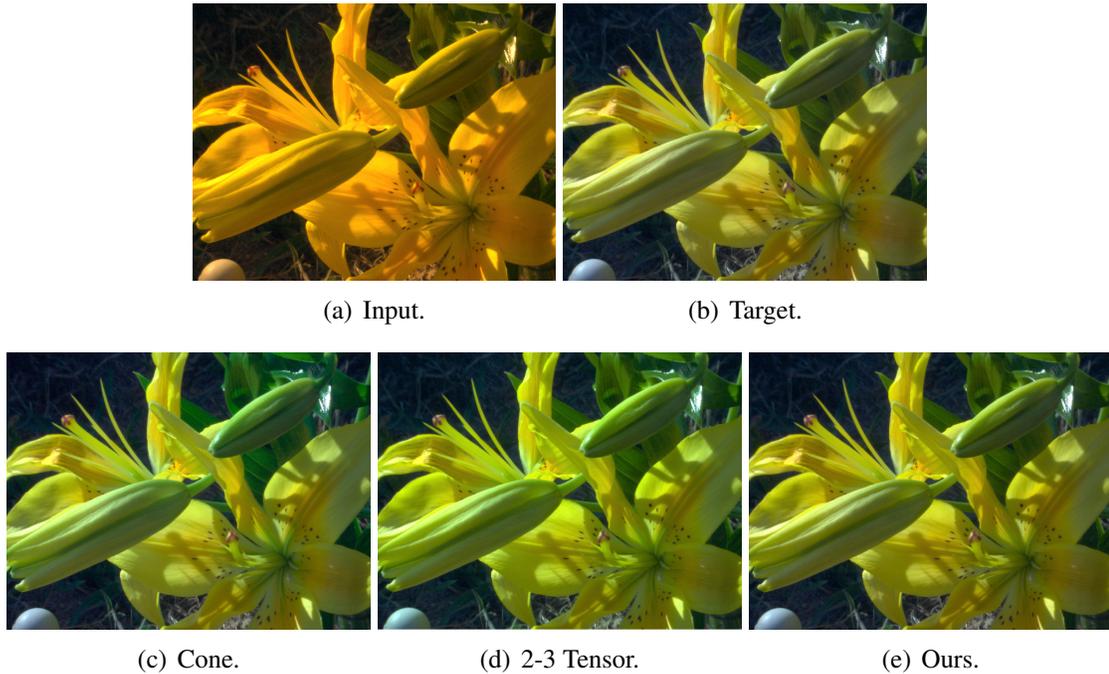


Figure 2.7: Color correction example. Diagonal mappings are used in three different color spaces to attempt a mapping of the input image (a) to the target illumination whose truth value is shown in (b).

We ran three different algorithms (Cone, 2-3 tensor, and Ours) on the SFU database to derive three optimized color bases for diagonal color constancy. Instead of choosing one particular algorithm for then computing diagonals, we wanted to characterize some notion of “best achievable performance” under each basis. We simulated the “best learned” diagonals with the following steps: transform the SFU derived measurement tensor by C^{-1} to obtain color coordinates in the candidate color basis; then for each material, compute the diagonal matrix mapping the color coordinates under the tungsten light bulb to the corresponding material color under daylight illumination (just component-wise ratios); finally, average the diagonal matrices over all materials to get the overall diagonal mapping between the two illuminants under the candidate color basis.

Figure 2.7 shows the output images for the three different color bases. The differences are subtle, but one can see that the cone basis yields leaves that are too green, and the 2-3 tensor method yields a flower that has too much blue (not enough of the red and yellow of the flower is present). Our method provides the best match to the ideal target image.

2.5.5 Constrained Optimization

For miscellaneous reasons, we may also want to constrain our optimized solution in other ways. For example, in Chapter 3 we seek a color basis in which the implied color gamut (collection of colors that have only positive components in the color basis) encompasses the entire set of human-visible colors. There, we also choose to measure perceptual error using the CIEDE2000 difference equation instead of using the standard ℓ_2 error in a linear color space. These extra constraints make each step of the alternating process of Section 2.3 a nonlinear least squares optimization. To handle the numerics, we used the Levenberg-Marquardt nonlinear least squares algorithm as implemented by [38] along with a wrapper function that enforces constraints.

The green dot in Figure Figure 2.4 represents the approximation error for the rank 3 tensor in which the color basis was constrained such that the implied color gamut would include the entire set of human-visible colors. If we use the green dot in place of the red dot for the rank 3 approximation, we still get a good approximation of the SFU tensor. The constrained rank 3 tensor's approximation error is 10.7% of the rank 2 tensor's approximation error. The rank 4 tensor's approximation error is 19.5% of the constrained

rank 3 tensor’s approximation error. We do not plot green dots for other tensor ranks because the “gamut constraint” is not so well-defined in those cases.

The color basis transform resulting from the constrained optimization is given by the following (which maps XYZ color vectors to the new color space):

$$C^{-1} = \begin{bmatrix} 9.465229 \times 10^{-1} & 2.946927 \times 10^{-1} & -1.313419 \times 10^{-1} \\ -1.179179 \times 10^{-1} & 9.929960 \times 10^{-1} & 7.371554 \times 10^{-3} \\ 9.230461 \times 10^{-2} & -4.645794 \times 10^{-2} & 9.946464 \times 10^{-1} \end{bmatrix} \quad (2.14)$$

See Figure 2.5 for the effective von Kries sensors, which are labeled as Alternating Non-linear Least Squares (ANLS).

In Chapter 3, we are particularly interested in the ability of the diagonal model to describe the effect of illumination change in the constrained color basis (whose gamut includes all the human-visible colors). Unlike the white patch normalization experiment which measured some notion of relational color constancy, we would like to directly measure the goodness of diagonal color constancy itself. To do this, we simulate a “perfect” white balancing algorithm that perfectly maps the color of a standard “white material” under one illuminant to the color of the same standard material seen under a different illuminant. In a von Kries compatible world, the same diagonal used to perform this mapping would correctly map all material colors under the initial illuminant to their corresponding colors under the second illuminant. We therefore apply the same mapping to each of the material colors under the first illuminant to derive predicted colors for the second illuminant. Deviations from the actual measured colors under the second illuminant then give a measure

| Median and RMS Diagonal Color Constancy Error for Constrained Color Basis | | | | | | | | |
|---|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | Tung. | Tung.+ | 3500K | 3500K+ | 4100K | 4100K+ | 4700K | 4700K+ |
| Tung. | | 1.67 2.28 | 0.54 0.84 | 2.23 2.98 | 1.14 1.54 | 2.72 3.63 | 1.74 2.31 | 3.20 4.34 |
| Tung.+ | 1.46 2.32 | | 1.07 1.69 | 0.54 0.80 | 0.60 0.97 | 1.02 1.40 | 0.31 0.55 | 1.54 2.11 |
| 3500K | 0.51 0.85 | 1.13 1.60 | | 1.68 2.24 | 0.57 0.79 | 2.16 2.92 | 1.19 1.59 | 2.67 3.65 |
| 3500K+ | 1.88 2.97 | 0.52 0.80 | 1.53 2.32 | | 1.07 1.67 | 0.58 0.77 | 0.66 1.09 | 1.14 1.55 |
| 4100K | 0.96 1.53 | 0.60 0.88 | 0.52 0.82 | 1.11 1.53 | | 1.60 2.18 | 0.61 0.84 | 2.10 2.89 |
| 4100K+ | 2.35 3.58 | 1.04 1.44 | 2.05 3.01 | 0.60 0.79 | 1.62 2.39 | | 1.14 1.71 | 0.58 0.81 |
| 4700K | 1.41 2.17 | 0.29 0.49 | 1.05 1.55 | 0.63 0.96 | 0.58 0.82 | 1.10 1.51 | | 1.59 2.17 |
| 4700K+ | 2.84 4.24 | 1.60 2.18 | 2.57 3.77 | 1.23 1.58 | 2.18 3.16 | 0.61 0.80 | 1.70 2.43 | |

Table 2.1: Results from the white balancing experiment. Top error is median error, bottom error is RMS error. Errors are measured in CIEDE2000 units. Row illuminant is the one that is chosen as canonical. The illuminants are taken from the SFU dataset: “Tung.” is a basic tungsten bulb (Sylvania 50MR16Q 12VDC); the different temperatures correspond to Solux lamps of the marked temperatures; the “+” means that a Roscolux 3202 Full Blue filter has been applied.

of the non-von Kries-ness of the world. This experiment also has the advantage of allowing error to be measured via the CIEDE2000 difference equation, giving some perceptual meaning to the resulting numbers. This also allows us to characterize performance in some absolute sense.

More specifically, we use the following procedure to report results. We choose some illuminant as a canonical illuminant. We also choose the one material in the SFU database labeled as “white” for our standard “white material.” Then, for every other (test) illuminant, we compute the mapping that perfectly maps the standard white under the

canonical illuminant to its color under the test illuminant and is diagonal in our constrained color basis. We then apply this mapping to all the material colors under the canonical illuminant to generate predictions for colors under the test illuminant. Errors between predictions and actual measured values in the SFU database are measured using the CIEDE2000 difference equation. This gives us, for each test illuminant, an error value associated with every material. So for each test illuminant, we report the median CIEDE2000 error and the root-mean-squared CIEDE2000 error of the material colors. See Figure 2.1 for the results.

2.6 Discussion

We have argued for a new data-driven choice of color basis for diagonal color constancy computations. We show that with respect to some existing metrics, the new choice leads to a better diagonal model.

While a linear change of color basis poses no problem to those concerned simply with algorithmic modeling, those who seek relevance to human biological mechanisms might object (on theoretical grounds) that sensor measurement acquisition may involve nonlinearities that disrupt the brain's ability to linearly transform the color basis downstream. Fortunately, experimental results based on single-cell responses and psychophysical sensitivity suggest that any existing nonlinearities at this level are negligible [41, 57].

Chapter 3

A New Color Space for Image Processing

In this chapter, we motivate the need for a new color space primarily by the desire to perform illumination-invariant image processing, in which algorithm outputs are not so sensitive to the illumination conditions of their input images. Simultaneously, we also seek a color space in which perceptual distances can be computed easily. We show that these desires relate very naturally to notions in perceptual science, and—with one additional assumption—fully lock down the form of the color space parameterization. We fit the remaining parameters to experimental data and apply our new color space to some examples to illustrate its utility.

3.1 Introduction

While existing color spaces address a range of needs, none of them simultaneously capture two notable properties required by a large class of applications that includes segmentation and Poisson image editing [48]. In this work, we present a new color space designed specifically to address this deficiency.

We propose the following two color space desiderata for image processing:

- (1) difference vectors between color pixels are unchanged by re-illumination;
- (2) the ℓ_2 norm of a difference vector matches the perceptual distance between the two colors.

The first objective restricts our attention to three-dimensional color space parameterizations in which color displacements, or gradients, can be computed simply as component-wise subtractions. Furthermore, it expresses the desire for these color displacements—the most common relational quantities between pixels in image processing—to be invariant to changes in the spectrum of the scene illuminant. Illumination invariance is useful for applications where processing is intended to operate on intrinsic scene properties instead of intensities observed under one particular illuminant. For example, Figure 3.3 (page 58) shows that when segmenting an image using usual color spaces, images that differ only in the scene’s illumination during capture can require nontrivial parameter tweaking before the resulting segmentations are clean and consistent. And even then, the segmentations are not necessarily reliable. Figure 3.4 (page 61) illustrates the extreme sensitivity a Poisson image editing algorithm exhibits when the illuminant of the foreground object does not

match the background's illumination.

The second condition implies that the standard computational method of measuring error and distance in color space should match the perceptual metric used by human viewers.

These desiderata have direct correspondence to widely-studied perceptual notions that possess some experimental support. Desideratum (1) corresponds to subtractive mechanisms in color image processing and to human color constancy. Desideratum (2) relates to the approximate “flatness” of perceptual space.

Subtractive mechanisms refer to the notion that humans perform spatial color comparisons by employing independent processing per channel [33, 43] and that such per-channel comparisons take a subtractive form. Physiological evidence for subtraction comes from experiments such as those revealing lateral inhibition in the retina [18] and the existence of double opponent cells in the visual cortex (where each type provides a spatially opponent mechanism for comparing a select chromatic channel) [46].

Color constancy is described in Section 1.1.1 and analyzed in detail in Chapter 2. Sometimes the term “chromatic adaptation” is used instead to emphasize the inability to achieve perfect constancy [18].

The approximate flatness of perceptual space refers to the relative empirical success—as evidenced by color spaces such as CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$ —in approximating perceptual distance with Euclidean distance when the Euclidean computation is preceded by an appropriate nonlinear reparameterization of color space.¹ There is ample

¹This property is called flatness since such compatibility with a Euclidean distance metric implies a lack

evidence that perceptual space is unlikely to be exactly flat [60], but there is also a fair amount of evidence that a large portion of it may be usefully treated as flat for many applications [31, 60].

Motivated by graphics applications, we explore the ramifications of enforcing these desiderata. The fact that both of our desiderata relate to human perceptual principles for low-level vision suggests that there is value in addressing each via an early-stage model, such as the choice of color space as opposed to the inclusion of additional structures at later stages of visual processing. We point out, however, that although the connections between our model and human perception are certainly suggestive, the utility of our approach does not depend on these assumptions regarding human visual processing being true.

To make our approach tractable, we make one further assumption.

- (3) materials and illuminants in our scenes are such that the effect of relighting is well-approximated by multiplying each tristimulus value (in an appropriate basis) by a scale factor that does not depend on the materials observed.

This assumption is commonly used in color correction algorithms [21]. As noted in Chapter 2, its veracity depends only on the spectral properties of the illuminants, the spectral reflectance of the materials, and the spectral sensitivities of the sensors (e.g., our cones). For evidence giving some quantitative support for this assumption, see Section 2.5.5. We emphasize that this assumption does not require any additional assumptions about human color processing.

Assumption (3) also relates to the (generalized) von Kries coefficient rule, which

of Riemannian curvature in the underlying manifold.

asserts that the brain is able to discount illumination changes by applying the appropriate inverse scale factors [18]. Color matching experiments evaluating the generalized von Kries hypothesis are therefore additional (indirect) tests of this assumption's veracity [60].

In this chapter, we show that conditions (1)-(3) almost completely determine the analytical form of the three-dimensional color space parameterization. There are only a handful of degrees of freedom in this model, and these can be “fit” using measurements provided by perceptual difference experiments.

Once our model is fit to perceptual data, we find that Euclidian distances in our color space better match a standard perceptual difference dataset than the widely used CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$ spaces.² Beyond its matching of perceptual data, the utility of our approach is supported by its apparent utility for image processing applications (Figure 3.3 and Section 3.5). In particular, we show that both segmentation and Poisson editing can be made robust to illumination changes simply by re-expressing colors in our color space prior to processing.

3.2 Previous work

There has been extensive research on formulating “uniform color spaces,” or color spaces in which perceptual distances correspond to Euclidean distances [31, 60]. The most widely used spaces of this type are the CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$ spaces. Some more recent

²In this regard, our method does not outperform some color difference equations such as CIEDE2000 and CIE94. As discussed in Section 3.4, however, such equations do not satisfy either of our desiderata, and this limits their applicability for many graphics applications.

attempts to predict perceptual distance modify CIE $L^*a^*b^*$'s Euclidean formula, but yield only difference equations, not parameterizations [40, 51]. A disjoint effort was carried out by researchers trying to determine a “line element,” or Riemannian metric, for color space [55, 60]. Despite leveraging the more general Riemannian formalism, a couple of the line elements nonetheless correspond to Euclidean spaces (e.g., Helmholtz, Stiles, Bouman-Walraven) [55]. The Riemannian metrics do not seem to be in wide use.

A large body of work also surrounds the human visual perception issues we touch upon. For more on the theories and their experimental assessments, see [46, 60] and the references therein. Significant research has also focused on computational aspects such as spatial methods for computing per-channel gain factors and estimating the illuminant [21, 33, 43]. Our approach is agnostic with regards to these methods and largely enables avoiding such computations altogether. Considerable work has also gone into investigating the conditions under which models such as the von Kries coefficient rule can work perfectly [14, 21, 58]. We build upon the results in Chapter 2 in section 3.3.3.

In this work, we pursue a functional equations approach to developing a new color space. This more directly incorporates empirical data into performing model selection in addition to using data for parameter estimation [11]. Perhaps the first to apply functional equations to psychophysics was Fechner, who proposed an explanation of Weber's Law—a law relating physical magnitudes of stimuli to perceived intensities [19]. His analysis, however, applies only for scalar judgments such as “brightness,” not for the space of colors more generally [39]. Mathematical psychologists have continued to refine the use of functional equations, but efforts seem mostly geared toward other areas of social science [1]. One exception is the work of Resnikoff, who derived two candidates for the structure of

color space—given a “homogeneity” assumption [49, 50]. One candidate, Resnikoff points out, is unlikely to be a model for actual color vision; the other is the Stiles line element. We start with somewhat different perceptual principles that have exhibited some quantitative success.

3.3 A Perceptual Metric

3.3.1 Formalizing the Color Space Conditions

Here we formalize the color space conditions presented in the introduction. Let us denote by \vec{x} the tristimulus values of a sensor represented in XYZ coordinates. Let us denote by F the 3D color space parameterization we wish to solve for.

According to our third assumption, when colors are written in some “appropriate” basis, the effect of relighting is simply a per-channel multiplication. We represent the fixed change to an appropriate basis by the matrix C^{-1} . (In Section 3.3.3, we show how to select an “optimal” matrix C^{-1} —and thus an optimal color basis—using a database of measured spectral distributions.) Under this assumption, the effect of relighting can be written as

$$\vec{x} \mapsto CDC^{-1}\vec{x}, \quad (3.1)$$

where D is a diagonal matrix depending only on the illuminants and not the materials.

Under our reparameterization F , then, color displacements transform as

$$F(\vec{x}) - F(\vec{x}') \mapsto F(CDC^{-1}\vec{x}) - F(CDC^{-1}\vec{x}'). \quad (3.2)$$

Desideratum (1) states that displacements are unchanged by relighting. Thus, for all diagonal matrices D , and color vectors \vec{x} and \vec{x}' we must have

$$F(\vec{x}) - F(\vec{x}') = F(CDC^{-1}\vec{x}) - F(CDC^{-1}\vec{x}'). \quad (3.3)$$

Desideratum (2) states that for any two colors \vec{x} and \vec{x}' , the perceptual distance between them is computed as

$$d(\vec{x}, \vec{x}') = \|F(\vec{x}) - F(\vec{x}')\|, \quad (3.4)$$

where $\|\cdot\|$ denotes the usual ℓ_2 -norm.

3.3.2 Form of the Metric

Given the assumptions in section 3.3.1, Appendix B.1 proves that the nonlinear function F of Equation (3.3) must take the form

$$F(\vec{x}) = A(\hat{\ln}(C^{-1}\vec{x})), \quad (3.5)$$

where A and C^{-1} are invertible 3×3 matrices and $\hat{\ln}$ denotes the component-wise natural logarithm. The matrix C^{-1} transforms color coordinates to the basis in which relighting (best) corresponds to multiplication by a diagonal matrix, while the matrix A provides degrees of freedom that can be used to match perceptual distances.

3.3.3 Parameter Estimation

We first optimize for the matrix C^{-1} using material and illuminant spectral data. Matrix A is then estimated using perceptual distance data.

Estimating the C^{-1} matrix

To compute the C^{-1} matrix, we adapt the approach in Chapter 2.

Let us consider the relighting assumption of Equation (3.1). This is necessarily an approximation, since given scenes with arbitrarily complicated material spectral reflection functions and arbitrarily complicated illuminant spectral densities, such a simple model—where the D matrix does not depend on the spectral properties of the material and C^{-1} is some pre-determined matrix mapping to a fixed preferred color basis—should not be possible. Recall from Chapter 2:

- Such a relighting model can in fact work if the collection of illuminants and materials contained in scenes has an associated measurement tensor that is at most rank 3.
- When the rank condition is satisfied by a collection of materials and lights, the C^{-1} matrix is fixed (up to 3 scale factors).

- A large collection of naturally-occurring materials and lights can be well-approximated while satisfying the rank condition.
- Given such a dataset, one can numerically solve for a locally-optimal C^{-1} matrix.

Since the functional form in Equation (3.5) involves a natural logarithm, we choose to use the constrained color basis whose gamut includes the entire set of human-visible colors. For more details see Chapter 2. The optimization procedure produced a rank three tensor approximation in which the colors had a root-mean-squared error of 1.1 CIEDE2000 units (a rough rule of thumb is that 1 CIEDE2000 unit is about 1 or 2 just noticeable differences). The mapping to our desired color basis for color constancy (assuming input coordinates in XYZ space) is given by

$$C^{-1} = \begin{bmatrix} 9.465229 \times 10^{-1} & 2.946927 \times 10^{-1} & -1.313419 \times 10^{-1} \\ -1.179179 \times 10^{-1} & 9.929960 \times 10^{-1} & 7.371554 \times 10^{-3} \\ 9.230461 \times 10^{-2} & -4.645794 \times 10^{-2} & 9.946464 \times 10^{-1} \end{bmatrix}. \quad (3.6)$$

To see the effective color sensors induced by using this color basis, see Figure 2.5.

Estimating the A matrix

We obtained two perceptual distance datasets for optimizing A . The RIT-Dupont dataset provides 312 pairs of colors that have been experimentally determined to be separated by a small and equal perceptual distance [8]. The Witt dataset provides both small and medium distance data for 418 color pairs [59].

Instead of optimizing for A directly, we first optimized for the matrix $G := A^T A$.

If for each color pair i , we define the displacement vector $\vec{v}_i := (\hat{\ln}(C^{-1} \vec{x}_i) - \hat{\ln}(C^{-1} \vec{x}'_i))$ and label its corresponding perceptual distance d_i , then we can rewrite Equation (3.4) as:

$$d_i(\vec{x}, \vec{x}') = (\vec{v}_i^T G \vec{v}_i)^{1/2}. \quad (3.7)$$

Since the \vec{v}_i and d_i are constants in the optimization, squaring Equation (3.7) leaves us with a set of constraints that are linear in the entries of G . Since G must be symmetric, we only have six variables to solve for. Such an overconstrained linear system is easily solved in a least-squares sense. As an aside, we also desire that G be positive definite; fortunately, our optimization naturally produced such an optimum, so no additional constraint was necessary.

To recover A from G , we must specify three more degrees of freedom (corresponding to rotations). We chose A to align with an eigenbasis for G . Computing an SVD of G gives us a decomposition of the form

$$G = U^T \Lambda U, \quad (3.8)$$

where U is a rotation matrix and Λ is diagonal. If we denote the diagonal matrix whose elements are the square roots of Λ 's by $\sqrt{\Lambda}$, we have

$$A = \sqrt{\Lambda} \cdot U. \quad (3.9)$$

For our purposes, we trained on the RIT-Dupont dataset and tested on both RIT-Dupont

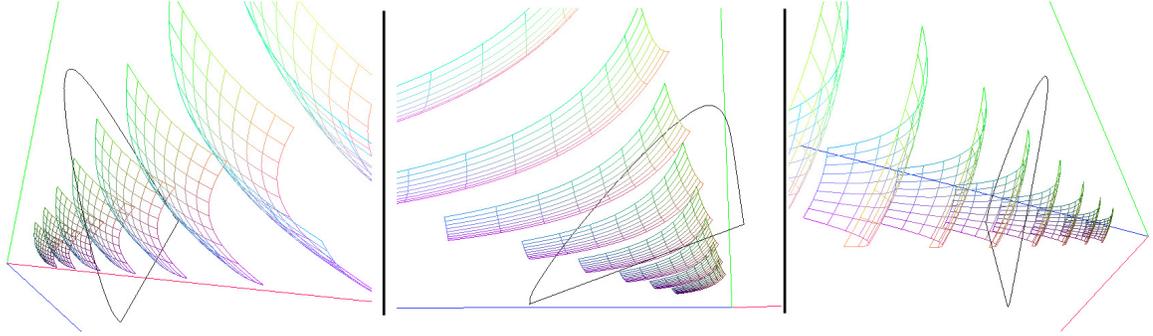


Figure 3.1: Iso-coordinate slices of a wireframe cube in our color space plotted in XYZ space under three different views. Lines connecting the slices of the cube were omitted to reduce clutter. The coordinate axes roughly correspond to “opponent channels.”

and Witt. The resulting matrix is given by

$$A = \begin{bmatrix} 2.707439 \times 10^1 & -2.280783 \times 10^1 & -1.806681 \\ -5.646736 & -7.722125 & 1.286503 \times 10^1 \\ -4.163133 & -4.579428 & -4.576049 \end{bmatrix}. \quad (3.10)$$

Figure 3.1 shows iso-coordinate slices of a wireframe cube in our color space (centered at a D65 whitepoint) visualized in XYZ space. Intervals between lines are all equal in our color space, but are nonlinearly spaced in the visualization since the mapping to XYZ space is nonlinear. To provide some landmarks, the cone of human-visible colors is drawn with a black line as it would appear on an xy-chromaticity diagram (it is projected onto a plane with constant $X + Y + Z$ value). The lines connecting the cube’s slices have been left out to reduce clutter in the diagram. These omitted lines are relatively straight in XYZ space, and roughly correspond to white-black directions in color space. The other two coordinate directions roughly correspond to blue-yellow and red-green directions. In this way, the coordinate system captures notions of “color opponency” [18, 46, 60].

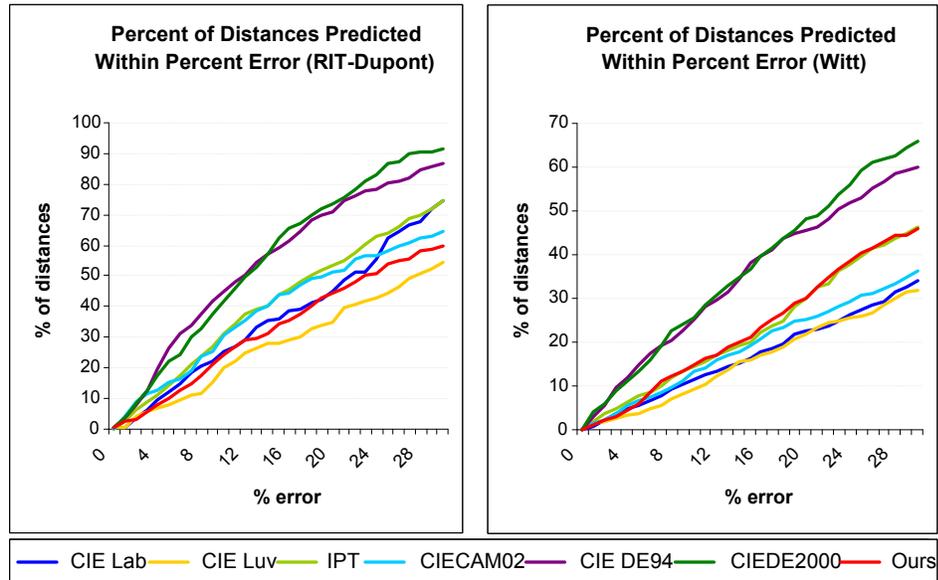


Figure 3.2: Perceptual error performance of our distance function against other color metrics. None of the other approaches satisfy all our desired requirements. For instance, CIEDE2000 and CIE94 do not provide 3D parameterizations, so it is not clear that they can be used in applications such as Poisson editing. Our metric was fit to the RIT-Dupont data.

3.4 Metric Results

In evaluating methods for computing perceptual distance, we were careful to not penalize uniform scalings of the distances. Rescaling a distance function does not change the color space geometry; it corresponds only to a different choice of units (e.g., although the numbers change depending on whether one uses inches or centimeters to measure spatial distance, such changes do not reflect *geometric* differences). For the RIT-Dupont dataset, in which all distances ought to be the same (without loss of generality, 1), we rescale computed distances such that the median distance is 1. For the Witt dataset, which has experimentally determined distances of varying magnitudes, we choose the rescaling that minimizes the ℓ_1 percentage error.

Figure 3.2 shows our metric’s performance on the RIT-Dupont and Witt datasets compared against other distance metrics. For RIT-Dupont’s small distance data, our metric performs comparably to CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$. For Witt’s small and medium distance data, it outperforms both. This suggests that our mapping (F) is more successful at mapping the geodesics of color space to straight lines.

The CIEDE2000 and CIE94 difference equations provide the best predictions of perceptual distance; however, these difference equations do not come with a 3D color space parameterization, so it is unclear whether they can be plugged into applications such as Poisson image editing. Even assuming these difference equations could be applied, the complexity of their formulas makes them less amenable to efficient computation, so many applications will likely benefit from simpler models that may sacrifice some of the accuracy. The work of [51] also points out that CIEDE2000’s fit comes at the cost of other desirable features such as the continuity of color space.

3.5 Applications

By transforming the tristimulus values of an image according to Equation (3.5), one obtains color descriptors that are (approximately) invariant to illumination. In this section we show that we can take advantage of these (approximately) illumination-invariant color descriptors to produce algorithms that operate on the “intrinsic image” instead of working directly on RGB values. We focus in particular on segmentation and Poisson editing, although the same ideas can be applied to other visual tasks such as feature tracking and texture synthesis.

3.5.1 Segmentation

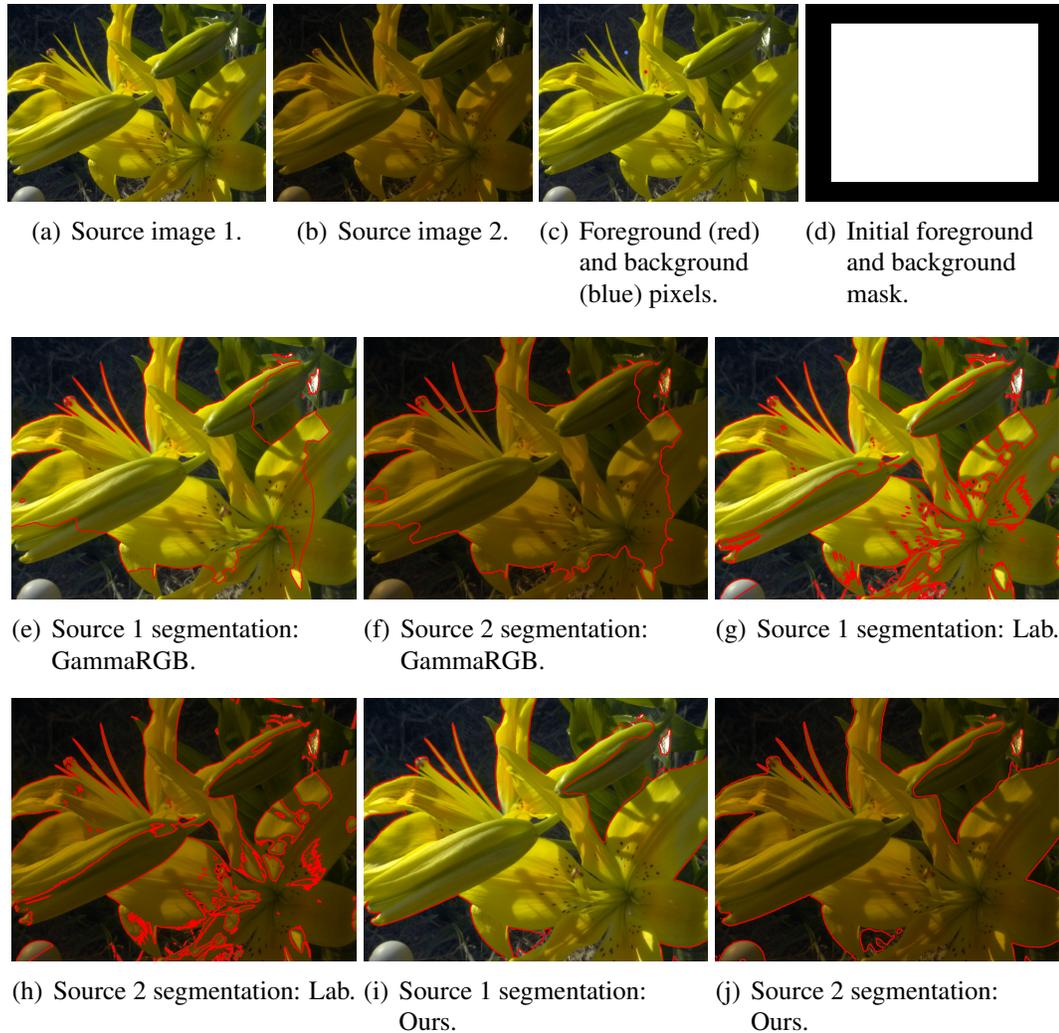


Figure 3.3: Segmentations of the same flower scene under two different illuminants. For (nonlinear) RGB and Lab color spaces, segmentation parameters for the two images were tweaked individually (optimal parameters for the image pairs were separated by 2 and 5 orders of magnitude respectively); for our color space, the same parameters were used for segmenting both images.

The objective of segmentation is often to separate a foreground object from background. If a segmentation algorithm is truly selecting “object,” then its segmentation should be invariant to changes in illuminant spectrum. If we assume that our world is (approx-

mately) consistent with our relighting assumption (expressed in Equation (3.1)), then color differences between neighboring pixels are invariant to illumination changes *when computation is done in our coordinate system*. Thus any segmentation algorithm that is based on color differences (gradients) between neighboring pixels will be (approximately) unaffected by global illumination changes.

We tested such illumination invariance using iterations of a geometric active contour algorithm as implemented by GAC++ [47]. Given a “hyperspectral image” (one possessing captured reflectance spectra for each pixel) from the dataset provided by [25], and daylight illuminants provided by [7], we rendered the hyperspectral image under two different illuminants to produce the input images shown in Figures 3.3(a) and 3.3(b). Since the segmentation algorithm required an estimated mean color for foreground and background, we chose arbitrary representative pixels for each (Figure 3.3(c)). We also provided the rough initial guess shown as a foreground/background mask (Figure 3.3(d)).

For each input image, we ran the segmentation algorithm in gamma-corrected RGB space (GammaRGB), CIE $L^*a^*b^*$ space (Lab), and our perceptual color space (Ours). The only free parameter to the segmentation algorithm is a constant controlling the discount factor for large gradients. For each run of the algorithm in gamma-corrected RGB space and CIE $L^*a^*b^*$ space, we manually tweaked the parameter to give the cleanest and most consistent segmentation between the light and dark image pairs. For the gamma-corrected RGB space, the optimal parameters were separated by two orders of magnitude. For the CIE $L^*a^*b^*$ space, the optimal parameters were separated by five orders of magnitude. When running the algorithm in our color space, we set the parameter once for one of the images, and ran the segmentation using the same parameter value for the other image (do-

ing so for the other spaces caused the algorithm to fail to even return an answer). The experiments showed that segmentation was relatively robust to illumination changes when run in our color space. Our hope is that such an approach, along with other methods, may help remove the need for manual tweaking of segmentation parameters.

3.5.2 Poisson Image Editing

Poisson image editing addresses the problem of deleting a portion of some image and filling it in using data from another source image [48]. We would like the algorithm to be insensitive to the illumination used for the inserted object. Once again, in our coordinates, color gradients are (approximately) unaffected by re-illumination. Thus, a Poisson algorithm which only uses the color gradients from its insertion source and works in our coordinates will be approximately invariant to the inserted element's illumination.

Let us denote the image being edited as I , and its damaged or masked out region by Ω . Let us denote the source image used to fill in Ω by J . In filling in Ω , our goal is to preserve the perceived color relations between the pixels in J as much as possible. This suggests that we solve the following problem [48]:

$$\min_I \int_{\Omega} \|\nabla I - \nabla J\|^2 dA, \quad \text{subject to } I \text{ being fixed on } \partial\Omega. \quad (3.11)$$

The Euler-Lagrange equations give us the following conditions:

$$\Delta I_i = \Delta J_i, \quad i = 1, \dots, \# \text{ color components.} \quad (3.12)$$



(a) Poisson experiment.

(b) Ours.



(c) Original image.

(d) LinearRGB.

(e) GammaRGB.



(f) CIECAM02.

(g) IPT.

(h) Lab.



(i) LogRGB.

(j) Ratio.

(k) Ours.

Figure 3.4: Three objects are inserted into a scene by running Poisson image editing in various color spaces.



(a) Poisson experiment.

(b) Ours.



(c) Original image.

(d) LinearRGB.

(e) GammaRGB.



(f) CIECAM02.

(g) IPT.

(h) Lab.



(i) LogRGB.

(j) Ratio.

(k) Ours.

Figure 3.5: A lizard is placed onto a tree by running Poisson editing in various color spaces.

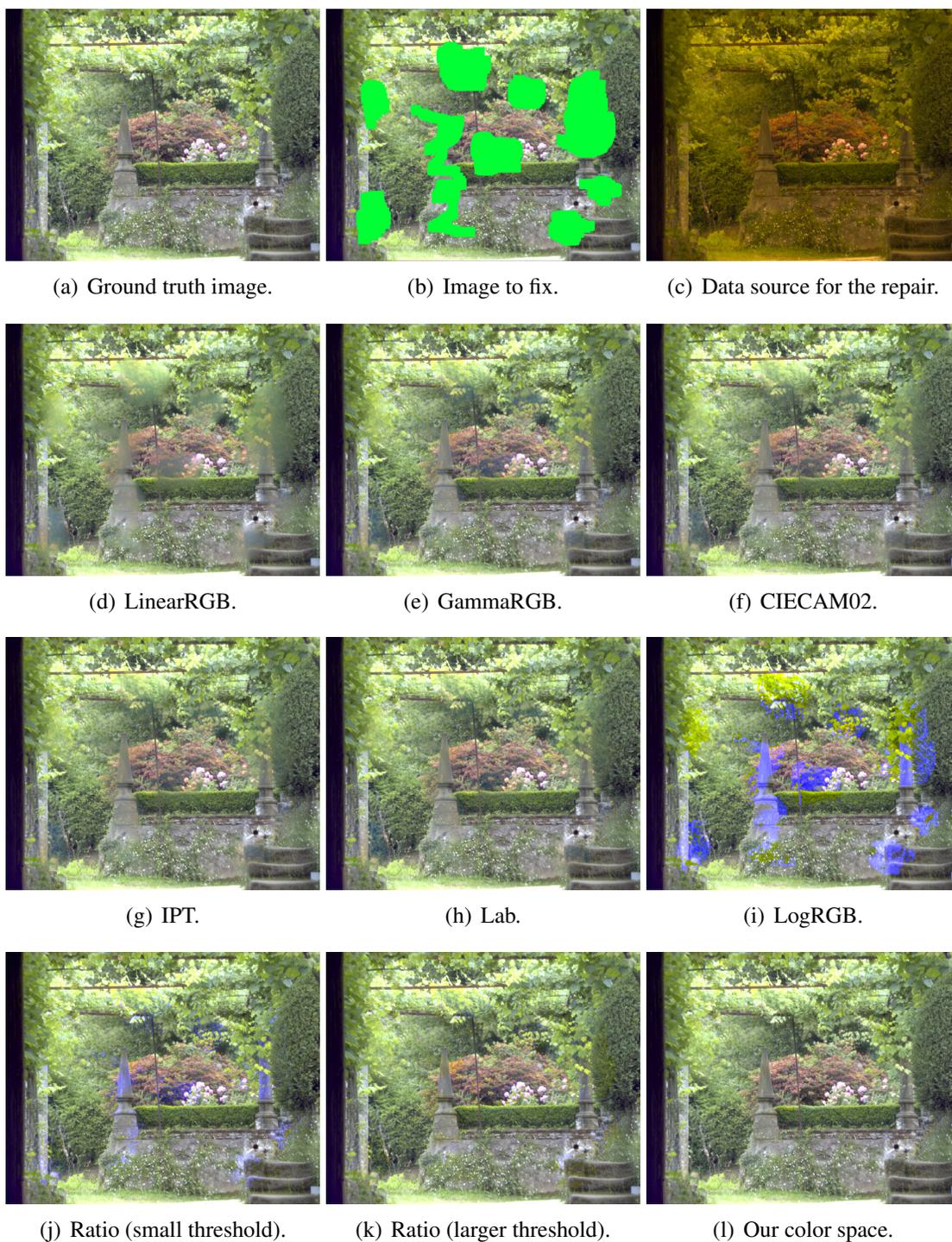


Figure 3.6: Filling in a damaged image using data from an image with different illumination conditions.

Figure 3.4 demonstrates a typical Poisson edit in which three objects that have been captured under different environmental conditions are inserted into a single scene. We show comparisons with results derived from solving Equation (3.12) in various other color spaces: gamma-corrected RGB (GammaRGB), linear RGB (LinearRGB), log of linear RGB (LogRGB), IPT [17], CIECAM02 [12, 44], and CIE $L^*a^*b^*$ (Lab).

We also compare against the related method of [27] (Ratio), which instead solves

$$\Delta \left(\frac{I_i}{J_i} \right) = 0, \quad i = 1, \dots, \# \text{ color components}, \quad (3.13)$$

where I and J are expressed in linear RGB coordinates.

To demonstrate performance in a case where ground truth is available, we also experiment with the (somewhat contrived) example in which the object to be inserted is actually coming from a version of the original image in its original location, the only difference being that this source image is captured under different illumination conditions.

The experiment is as follows: render a hyperspectral image under two different illuminants (as in Section 3.5.1); in one image (denoted by I above), excise data in some region (denoted by Ω); use the other image (denoted by J) along with I 's excision mask to provide source data for filling in Ω . The output of each algorithm can then be compared against the original image I before the data excision. A perfect algorithm is one that produces exact matches.

We ran the experiment for a couple different hyperspectral images and different pairs of illuminant spectra. Figure 3.6 shows one of the runs. These images have been

designed for display on a monitor calibrated with a gamma of 2.2 and D65 white point. Ratio and LogRGB were both able to handle mild forms of illumination change because they account for multiplicative scalings of linear RGB. However, their implicit use of the RGB basis for von Kries mappings and failure to use perceptual weightings on the RGB channels led to some noticeable artifacts for more extreme illumination changes. In Figure 3.6, we see both suffer from taking the logarithm or dividing by a small value. Some of the problems in Ratio were ameliorated by raising the min-value threshold (for algorithm details, see [27]). The threshold lower bounds the darks in J , and when set too high causes noticeable chromaticity shifts (in particular, note the discontinuous green chromaticity shift on the bush on the right in Figure 3.6(k) and the excessive yellowness added to the flowers on left). In Figure 3.6, the threshold for the two Ratio images correspond to 10 gamma-corrected RGB units and 20 units respectively. Lab generally outperformed GammaRGB. LinearRGB always performed worst. IPT and CIECAM02 performed very similarly to CIE $L^*a^*b^*$ in all our experiments, so we omit showing their outputs in Figure 3.6.

3.6 Discussion

We have presented a new color space that endows image processing algorithms with some robustness to illumination change. The associated metric also matches the Witt perceptual difference dataset better than CIE $L^*a^*b^*$ and CIE $L^*u^*v^*$ (and performs comparably on RIT-Dupont), but is equally easy to use. We have also drawn some connections between our desiderata and various human perceptual principles. It would be interesting to look for further perceptual correspondences. For instance, one wonders to what extent the deviation

from perfect color constancy in humans can be predicted by the use of our color space.

With regards to applications, continued acquisition of color constancy and perceptual difference datasets (perhaps with more saturated colors as well), and continued search for even better optima based on these datasets, may further improve our metric's predictive ability. We note that if only illumination invariance is desired, much of the invariance comes from the combination of using logarithms and choosing a C^{-1} matrix such that the implied color gamut tightly includes the full range of human visible colors (larger than typical RGB working space gamuts but smaller than the implied cone gamut).

Chapter 4

Differential Geometry

In this chapter, we lay the mathematical foundations we will need for discussing the relationships between curves and surfaces. The techniques we discuss come from the subfield of mathematics known as differential geometry. Our coverage here will likely be too concise to serve as an initial introduction to the material; we instead outsource more thorough treatments to some standard references: for a more concrete and classical approach, the book by Gallier provides a nice introduction [26]; for thorough grounding in curves and surfaces, the book by Do Carmo is considered a standard text [9]; for more modern notation, we refer the reader to the first couple chapters of the General Relativity text by Wald [56]—Wald is also extremely economical in his presentation, so the text by Carroll may be used to fill in some of the exposition [10]; for an encyclopedic reference, consult [53]. Unfortunately, each text employs different notation and makes combining lessons learned from each a nontrivial task (e.g., modern notation, classical interest in curves and surfaces, noncoordinate versus coordinate bases, etc.). Therefore, this exposition is in-

tended to do the following: (1) introduce notation we will be using throughout (we adopt the modern abstract index notation), and (2) emphasize aspects relevant for our use (we will be concerned with both intrinsic and extrinsic information and would like to work in a coordinate-independent manner).

4.1 Manifolds, Tensors, and Calculus

Roughly speaking, a manifold can be thought of as a continuum of fixed dimension (we will make this intuition much more precise in section 4.1.1). Manifolds enjoy frequent use in mathematics because they provide a generalized setting in which the rules of calculus can still be applied. They also prove extremely useful in describing the laws of physics and hence work their way into many applications.

When studying shape, we are concerned with smooth surfaces embedded in \mathbb{R}^3 . Both the smooth surface and the ambient space \mathbb{R}^3 are modeled as manifolds. When studying color, the space of colors is also modeled as a manifold. Crucially, in the case of color, the manifold does not come with a natural embedding into a Euclidean space, so an understanding of both intrinsic and extrinsic descriptions of manifolds is useful. Even in the case of surfaces where a natural embedding into \mathbb{R}^3 is associated with the surface, distinguishing between intrinsic and extrinsic information can prove useful.

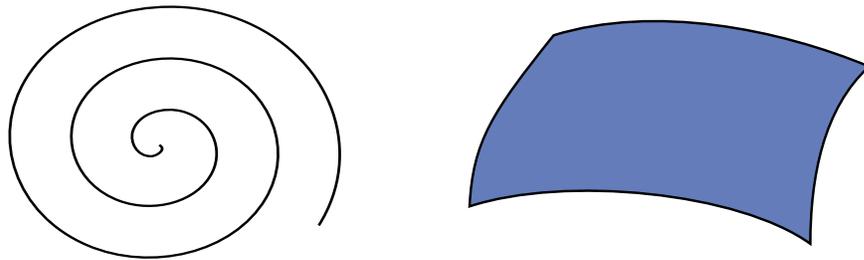


Figure 4.1: Manifolds.

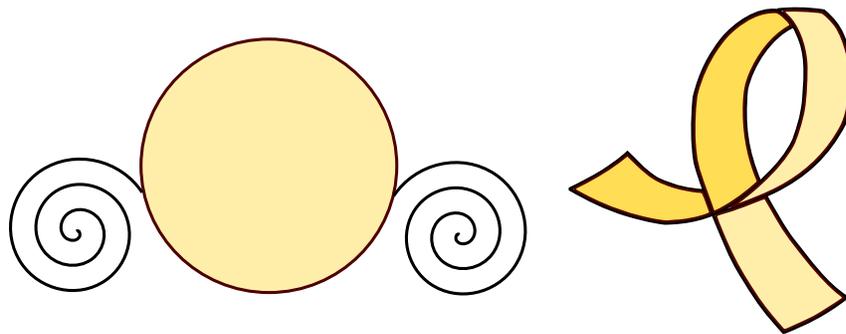


Figure 4.2: Non-manifolds. Points joining pieces with different dimensions and points of self-intersection are not allowed for manifolds.

4.1.1 Manifolds

A manifold is a set where every local patch “looks like” an open subset of Euclidean space. The patches must also be “stitched together” in a consistent manner. Figure 4.1 depicts two examples. Figure 4.2 depicts two non-manifolds for comparison. Formally, we say that M is a *topological manifold of dimension n* if it is a paracompact Hausdorff space that is locally homeomorphic to \mathbb{R}^n . Hausdorff-ness and paracompactness are technical conditions that disallow pathological objects from consideration and allow us to define integration on the resulting object. “Locally homeomorphic to Euclidean space” means

there is a collection of subsets $\{U_\alpha\}$ such that:

- (1) Every point p in M belongs to at least one subset U_α .
- (2) For each U_α , there is a homeomorphism $x_\alpha : U_\alpha \rightarrow V_\alpha$, where V_α is an open subset of \mathbb{R}^n .

The homeomorphisms $\{x_\alpha\}$ are called (*coordinate charts*), and the homeomorphisms $x_\beta \circ x_\alpha^{-1} : x_\alpha(U_\alpha \cap U_\beta) \rightarrow x_\beta(U_\alpha \cap U_\beta)$ (defined for every U_α and U_β with non-empty intersection) are called *transition maps*.

A *differentiable manifold* additionally possesses *differential structure* (i.e., the transition maps are not only continuous, but smooth); this bars the existence of kinks and creases in the manifold. In dimensions less than four, every topological manifold admits a unique differential structure, so we do not have to worry about any ambiguity in this sense. For a more detailed discussion, see [30].

4.1.2 Tensor Fields

On manifolds, we can define various objects, such as functions, vectors, covectors, and higher order tensors. Often, simply keeping track of the variable types and their various inter-relations can lead to notational nightmares. To minimize confusion, we will adhere to abstract index notation. Abstract index notation has the advantage of allowing us to work in a coordinate independent manner; and when basis-specific computations become necessary, we have the flexibility of choosing any basis we like (even “non-coordinate” ones).

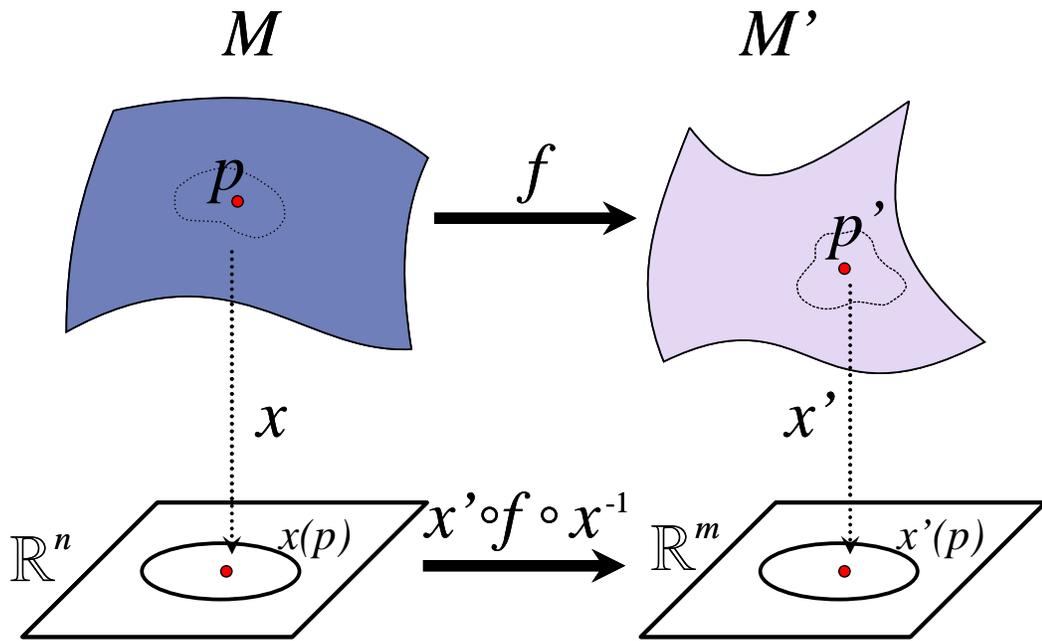


Figure 4.3: A mapping between manifolds induces a mapping between parameter domains.

Functions

Functions assign a real number to every point p in an n -dimensional manifold M . We have already seen examples of these: the coordinate charts, $\{x_\alpha\}$. Each coordinate chart is actually a collection of n functions. The μ^{th} function maps a point p in M to its μ^{th} coordinate value. To denote the μ^{th} coordinate value, it is common to use μ as a superscript: $x_\alpha^\mu(p)$. It is also common to suppress the α subscript (marking the particular subset of M used to define a homeomorphism with Euclidean space) and simply write $x^\mu(p)$.

We can also define functions between manifolds (see Figure 4.3). Such a function f maps a point $p \in M$ to a point $p' \in M'$. Given charts $\{x_\alpha\}$ for M and $\{x'_\beta\}$ for M' , we say that f is smooth if $x'_\beta \circ f \circ x_\alpha^{-1}$ (which is a map between Euclidean spaces) is smooth (whenever such a construction makes sense).

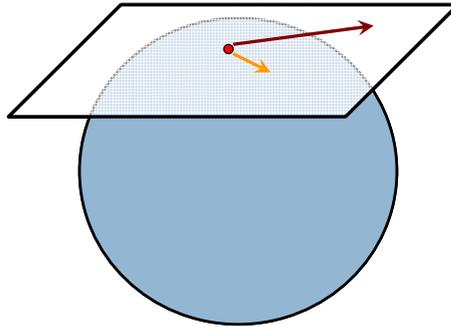


Figure 4.4: Two vectors tangent to the manifold (sphere in this case).

Vectors

Vectors are of course a ubiquitous and useful concept in geometry. In differential geometry, vectors always live in the space tangent to a manifold and go by a lot of different names. They may be referred to as “tangent vectors,” “contravariant vectors,” or simply as “vectors.” We will just call them vectors. Figure 4.4 illustrates the intuitive concept.

For example, in shape analysis we have two manifolds, the ambient space \mathbb{R}^3 and the embedded surface. Each of these has a tangent space defined at each point. For \mathbb{R}^3 , the tangent space at any point is just \mathbb{R}^3 itself, and so vectors on this manifold are just the familiar 3-vectors (pictorially, we may think of these 3-vectors as being translated so as to be anchored at the point of interest). As illustrated in Figure 4.4, vectors on the embedded surface are always restricted to lie in the plane tangent to the surface at the point of interest. Since the surface is embedded in \mathbb{R}^3 , these surface vectors naturally correspond to vectors in \mathbb{R}^3 as well.

In standard vector notation, we might represent a vector \vec{v} in \mathbb{R}^3 by first picking

some basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ for \mathbb{R}^3 and then writing the 3-vector as

$$\vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3 \quad (4.1)$$

$$= \sum_{\mu=1}^3 v_\mu \vec{e}_\mu. \quad (4.2)$$

To reduce notational clutter, we adopt the Einstein summation convention in which repeated indices, with one raised and one lowered, denote summation over the possible index values. As a convention, the components of vectors are given raised indices, and vectors themselves have lowered indices. So we rewrite the above expression as

$$\vec{v} = v^1 \vec{e}_1 + v^2 \vec{e}_2 + v^3 \vec{e}_3 \quad (4.3)$$

$$= v^\mu \vec{e}_\mu. \quad (4.4)$$

In abstract index notation, we furthermore do away with the arrow denoting certain objects as vectors, and use placeholder indices instead. We use Latin indices (e.g., a, b, c, \dots) to denote placeholders that signify the data-type of the tensor. We will reserve Greek indices (e.g., μ, ν, σ, \dots) for basis specific entries such as coordinate components.¹ So Greek indices only make sense after we have chosen a particular basis. The convention for vectors is to include the Latin placeholder as a superscript. With these rules, we rewrite the above expression again to obtain

$$v^a = v^\mu (e_\mu)^a. \quad (4.5)$$

¹When the use of Greek letters may lead to confusion, we will instead use Latin indices starting from “ i ” (e.g., i, j, k, ℓ, \dots).

As noted above, vectors always live in the space tangent to the manifold. This is because only such objects can be defined without reference to the way the manifold is embedded in some larger space (i.e., are intrinsic). The formalism for such an intrinsic definition relies on identifying tangent vectors with directional derivatives. Therefore, we can think of any vector as acting on functions by taking directional derivatives (geometrically, the vector points in the direction of interest). Notationally, we'll write the action of a vector v^a on a function f as

$$v[f]. \quad (4.6)$$

Given a coordinate chart x^μ , the directional derivative can be written as

$$v[f] = v^\mu \frac{\partial (f \circ x^{-1})}{\partial x^\mu}, \quad (4.7)$$

where on the left-hand side, the vector v^a is defined for points p on the manifold; and on the right hand side, the quantities are defined for each corresponding point $x(p)$ in the range of the chart. Because of this identification of vectors with directional derivatives, the basis vectors (for the tangent space) associated with the coordinate chart are written as $\left(\frac{\partial}{\partial x^\mu}\right)^a$.

Given two coordinate charts x^μ and $x'^{\mu'}$ for an open set around a point p on a manifold, a vector v^a defined at p has coordinates v^μ and $v'^{\mu'}$ with respect to the two coordinate charts. The coordinate functions for the vector are related by

$$v^\mu = v'^{\sigma'} \frac{\partial x^\mu}{\partial x'^{\sigma'}}. \quad (4.8)$$

This transformation law for vectors essentially comes from the chain rule of multivariable calculus (one treats vectors as directional derivatives and mandates their action on functions is independent of the coordinate chart). This form of transformation law for coefficients is sometimes called “contravariant” and is where some texts derive the convention of calling vectors “contravariant vectors.” We explicitly avoid using such terminology because vectors themselves may be thought of as “covariant.” Naming objects based on whether they are thought of in terms of their components or their entirety seems to be an undeserved source of confusion. Henceforth, we will not use the term “contravariant.” The adjective “covariant” will only show up in the context of *covariant derivatives*, in which it has a completely different meaning.

Covectors

A covector is a linear functional on vectors. We mark a covector as such by giving its variable name a Latin subscript (e.g., w_a). Given a vector v^a , we can denote the action of a covector w_a on v^a (or contraction of the two) by placing them side-by-side (order does not matter) and making them share the same Latin index: $w_a v^a$. Notationally, this looks similar to the Einstein summation convention, but there is no summing over indices here.

Given an n -dimensional vector space, we can choose some basis, $\{(e_\mu)^a\}_{\mu=1}^n$, for that space. We can also define a dual basis $\{(\vartheta^v)_a\}_{v=1}^n$ such that

$$(e_\mu)^a (\vartheta^v)_a = \delta_\mu^v, \quad (4.9)$$

where δ_μ^v is 1 when μ equals v and 0 otherwise. Note that this duality depended on

the choice of basis we made for vectors; there is no “canonical” isomorphism between the spaces unless we have additional information like a preferred set of basis vectors or a metric (which will be described later).

A general covector, being a linear functional on vectors, must live in the linear span of the $\{(\vartheta^V)_a\}$. We can then give a covector w_a coordinates w_V with respect to this dual basis. Using the Einstein summation convention, we have the relation $w_a = w_V(\vartheta^V)_a$ for some set of coordinates $\{w_V\}$.

Given two coordinate charts x^μ and $x'^{\mu'}$ for an open set around a point p on a manifold, a covector w_a defined at p has coordinates w_μ and $w'_{\mu'}$ with respect to the two coordinate charts. The coordinate functions for the covector are related by

$$w_\mu = w'_{\sigma'} \frac{\partial x'^{\sigma'}}{\partial x^\mu}. \quad (4.10)$$

Higher Order Tensors

Many interesting quantities depend on vector displacements in a multilinear (linear in each variable) fashion (e.g., fluid stress, linear elasticity, heat flux). These quantities, called *tensors*, are said to have geometrical, physical, or coordinate-independent meaning because given tensor measurements in one particular reference frame, it can easily be determined what measurements would have resulted in any other reference frame.

More generally, we are interested in *tensor fields*, where a different tensor is attached to each spatial location² (i.e., point on the manifold). The spatial transition between

²The use of the word ‘spatial’ here refers to a location on the manifold. It does not imply that such

tensors is assumed to be smooth. In practice, it is common to define tensor fields but refer to them as tensors, with the implicit understanding that the tensor fields are being evaluated at a particular spatial point of interest. The set of tensor fields turns out to be a rich enough class of objects to describe just about any quantity of interest.

One tensor (field) of particular interest is the metric tensor g_{ab} which acts on two vectors v^a and w^b and yields their inner product. The metric tensor appears so often that it is always given the name “ g ”. Since g_{ab} is multilinear, it can be expanded in a basis $\{(\vartheta^\mu)_a \otimes (\vartheta^\nu)_b\}$ where $\{(\vartheta^\mu)_a\}$ is a basis for covectors and $(\vartheta^\mu)_a \otimes (\vartheta^\nu)_b$ acts on the ordered pair of vectors (v^a, w^b) by

$$((\vartheta^\mu)_a \otimes (\vartheta^\nu)_b)(v^a, w^b) := ((\vartheta^\mu)_a v^a) \cdot ((\vartheta^\nu)_b w^b). \quad (4.11)$$

To reduce the notational clutter, it is common to omit the ‘ \otimes ’ symbol. The Latin superscripts and subscripts maintain clarity of what’s going on. In coordinates, the metric tensor can be written as

$$g_{ab} = g_{\mu\nu} (\vartheta^\mu)_a (\vartheta^\nu)_b. \quad (4.12)$$

The introduction of a metric tensor also induces a preferred isomorphism between vectors and covectors. In particular, given a vector v^a , we map it to its corresponding covector v_a using the metric:

$$v_a = g_{ab} v^b \quad (4.13)$$

positioning on the manifold corresponds to positions in the common 3-space that we explore.

We have chosen to use the same variable name for the vector and covector since the Latin indices automatically note that they differ only via the duality coming from the metric. Since g_{ab} is symmetric, it does not matter which index of the metric tensor we contract along. We can also map covectors to vectors by employing the inverse metric tensor (field), g^{ab} . Although we use the same name “g” for the inverse metric tensor as was used for the metric tensor, this inverse metric tensor is a completely different tensor. This is the one case where, because of convention, we do this. In coordinates, the inverse metric tensor is written as:

$$g^{ab} = g^{\mu\nu} (e_\mu)^a (e_\nu)^b \quad (4.14)$$

The components for the metric tensor and the inverse metric tensor are related by:

$$g_{\mu\nu} g^{\nu\sigma} := \delta_\mu^\sigma \quad (4.15)$$

A covector w_a is then mapped to a vector w^a by

$$w^a = g^{ab} w_b. \quad (4.16)$$

The duality between vectors and covectors given by a metric also gives meaning to general tensor fields with any number of upper and lower indices. For example a tensor field T^{abc}_{de} , when evaluated at a point on the manifold, yields a tensor that acts on three covectors and two vectors to produce a number. One can also interpret the tensor as acting on two covectors and two vectors to produce a vector. Likewise for other pairings of

arguments and indices. The *type* of a tensor is (p, q) if it has p upper indices and q lower indices.

Given two coordinate charts x^μ and $x'^{\mu'}$ for an open set around a point p on a manifold, a tensor $T^{a_1 \dots a_r}_{b_1 \dots b_s}$ defined at p has coordinates $T^{\mu_1 \dots \mu_r}_{\nu_1 \dots \nu_s}$ and $T'^{\mu'_1 \dots \mu'_r}_{\nu'_1 \dots \nu'_s}$ with respect to the two coordinate charts. The coordinate functions for the tensor are related by the *tensor transformation law*:

$$T^{\mu_1 \dots \mu_r}_{\nu_1 \dots \nu_s} = T'^{\sigma'_1 \dots \sigma'_r}_{\lambda'_1 \dots \lambda'_s} \frac{\partial x^{\mu_1}}{\partial x'^{\sigma'_1}} \cdots \frac{\partial x^{\mu_r}}{\partial x'^{\sigma'_r}} \frac{\partial x'^{\lambda'_1}}{\partial x^{\nu_1}} \cdots \frac{\partial x'^{\lambda'_s}}{\partial x^{\nu_s}}. \quad (4.17)$$

It is perhaps not obvious at this point that there exist choices of local bases, called noncoordinate bases, which do not arise from any coordinate chart. One particular choice—which we will make repeated use of—is an orthonormal frame field (we use the terms “frame” and “basis” interchangeably). The fact that these bases do not generally correspond to coordinate bases is discussed in Appendix C.1. So to generalize Equation (4.17), we note that the tensor component transformations are really about basis changes, not just coordinate chart transformations. An analogous expression holds for noncoordinate bases where the Jacobians are simply replaced by matrices representing the basis changes.

4.1.3 Integrals

The general power of differential geometry comes from the fact that it is possible to do calculus on manifolds. Here we revisit the notion of integration, which can be defined on a (differentiable) manifold even if it has no metric structure.

Differential forms

To define integration, we first define the notion of differential forms. The prototype for differential forms is the determinant function of standard vector calculus that, given coordinates of vectors, computes parallelepiped areas and volumes.

Differential forms take these ideas, but express them in a coordinate-free formulation. The defining properties of the determinant are its multilinearity and antisymmetry. Differential p -forms are therefore defined to be totally antisymmetric $(0, p)$ tensors. In other words, $T_{a_1 \dots a_p}$ is a differential p -form if:

$$T_{a_1 \dots a_p} = T_{[a_1 \dots a_p]}, \quad (4.18)$$

where the right-hand side is the antisymmetrization of the tensor. Antisymmetrization of a $(0, p)$ tensor can be written more explicitly as

$$T_{[a_1 \dots a_p]} := \frac{1}{p!} \sum_{\pi} \text{sgn}(\pi) T_{a_{\pi(1)} \dots a_{\pi(p)}}, \quad (4.19)$$

where π is a permutation and $\text{sgn}(\pi)$ is 1 if the permutation is even and -1 if it is odd. As a side remark, a 0-form is just a function over the surface.

There is also a map, \wedge , called the wedge-product that maps p -forms and q -forms to $(p + q)$ -forms:

$$v_{a_1 \dots a_p} \wedge w_{b_1 \dots b_q} := \frac{(p + q)!}{p!q!} v_{[a_1 \dots a_p} w_{b_1 \dots b_q]}. \quad (4.20)$$

So for example, the wedge-product between two 1-forms v_a and w_b is given by

$$v_a \wedge w_b = v_a w_b - v_b w_a. \quad (4.21)$$

If $(p + q)$ is larger than the dimension of the manifold, the resulting differential form is zero. The wedge-product of a function (0-form) and a p -form is simply the multiplication of the p -form by the function.

Integration

We now move to discussing integration over orientable manifolds. It is common to picture the orientation of a surface as a consistent choice of normal (say, outward pointing) over the entire surface. Such a definition cannot be used in general, however, since it uses information extrinsic to the manifold. We can, instead, use differential forms to define orientation. Over an n dimensional manifold M , the space of n -forms is 1 dimensional. If $\epsilon_{a_1 \dots a_n}$ is a non-vanishing n -form on the manifold, we say that $\epsilon_{a_1 \dots a_n}$ defines an orientation. Two orientations are said to be equivalent if one n -form is simply an everywhere positive scaling of the other. Under a coordinate chart x^μ for an open subset U of M , the orientation n -form can be written as

$$\epsilon_{a_1 \dots a_n} = h \cdot (dx^1)_{a_1} \wedge \dots \wedge (dx^n)_{a_n}. \quad (4.22)$$

In a right-handed coordinate system, h is positive function. (For convenience, we will always choose a right-handed orientation.) Any other n -form $w_{a_1 \dots a_n}$ can be written as:

$$w_{a_1 \dots a_n} = f \cdot (dx^1)_{a_1} \wedge \dots \wedge (dx^n)_{a_n}. \quad (4.23)$$

The integration of $w_{a_1 \dots a_n}$ with respect to the orientation $\varepsilon_{a_1 \dots a_n}$ is then defined to be

$$\int_U w_{a_1 \dots a_n} := \int_{x(U)} f dx^1 \dots dx^n, \quad (4.24)$$

where the right-hand side is standard Lebesgue integration in \mathbb{R}^n . If we used a different (orientation-preserving) coordinate chart x'^{μ} , our n -form $w_{a_1 \dots a_n}$ would be written as

$$w_{a_1 \dots a_n} = f' \cdot (dx'^1)_{a_1} \wedge \dots \wedge (dx'^n)_{a_n}. \quad (4.25)$$

The tensor transformation law for tensors tells us that

$$f' = f \det \left(\frac{\partial x^\mu}{\partial x'^{\nu'}} \right). \quad (4.26)$$

This replicates the standard rule for coordinate changes of integrals in \mathbb{R}^n , so our definition is coordinate chart-independent. The paracompactness of our manifold then allows us to extend this definition of integration over a patch to integration over the entire manifold. “Pullbacks” of differential forms onto embedded submanifolds then also gives meaning to integration of p -forms over p dimensional embedded submanifolds (see section 4.2.3 for more on pullbacks).

The presence of a metric furthermore gives us a natural choice of volume element for integration. We choose the n -form $\varepsilon_{a_1 \dots a_n}$ such that

$$\varepsilon^{a_1 \dots a_n} \varepsilon_{a_1 \dots a_n} = n!. \quad (4.27)$$

In a right-handed coordinate system, this yields the natural volume element:

$$\varepsilon_{a_1 \dots a_n} = \sqrt{|g|} (dx^1)_{a_1} \wedge \dots \wedge (dx^n)_{a_n}, \quad (4.28)$$

where $|g|$ denotes $\det(g_{\mu\nu})$.

4.1.4 Derivatives

The other notion necessary for calculus is the notion of derivative. The directional derivative of a scalar function, defined earlier, is an intuitive one. However, generalizing this notion to higher order tensors is less straightforward. In fact, in differential geometry, there is a plethora of derivatives, each acting the same on functions but acting differently on higher order tensors.

Partial derivatives

Perhaps the easiest derivative operator to define is the partial derivative, ∂_a . The use of a Latin subscript on ∂_a is a commonly employed notational convenience even though the partial derivative is not itself a covector. The lower index simply marks that it acts on a type (r, s) tensor to produce a type $(r, s + 1)$ tensor. Given a coordinate chart for a manifold, the

partial derivative associated with that chart acts on a tensor $T^{b_1 \dots b_r}_{c_1 \dots c_s}$ with coordinates $T^{\sigma_1 \dots \sigma_r}_{\rho_1 \dots \rho_s}$ by

$$\partial_a T^{b_1 \dots b_r}_{c_1 \dots c_s} := (dx^\mu)_a \frac{\partial}{\partial x^\mu} [T^{\sigma_1 \dots \sigma_r}_{\rho_1 \dots \rho_s}] \left(\frac{\partial}{\partial x^{\sigma_1}} \right)^{b_1} \cdots \left(\frac{\partial}{\partial x^{\sigma_r}} \right)^{b_r} (dx^{\rho_1})_{c_1} \cdots (dx^{\rho_s})_{c_s}. \quad (4.29)$$

Partial derivatives are natural in the sense that we know there exists a set of charts that cover any particular manifold, and the mapping to Euclidean space allows us to use the derivatives we're familiar with in \mathbb{R}^n . However, the definition of partial derivatives is dependent on the choice of chart. Therefore, if we define any quantity using partial derivatives (of anything higher order than a function), the definition of that object would in general depend on the choice of coordinates. Although such coordinate dependence is permissible in certain contexts, it is often desirable to perform coordinate-independent manipulations. After all, the actual geometry does not depend on the coordinate frames we use to describe it.

Exterior derivatives

One coordinate independent derivative that can be defined without any additional structure on a (differentiable) manifold is the exterior derivative, d_a . Again, we use the lower Latin index to note that it maps a $(0, p)$ type tensor to a $(0, p + 1)$ tensor. In fact, it is only applicable on an even smaller subset of tensors: the differential forms. Since d_a maps a r -form to a $(r + 1)$ -form, it is defined using antisymmetrization:

$$d_a w_{b_1 \dots b_r} := \partial_{[a} w_{b_1 \dots b_r]}. \quad (4.30)$$

Even though the definition of d_a involved the use of partial derivatives, ∂_a , the antisymmetrization kills all terms that possess coordinate chart dependence. One important property of the exterior derivative is that applying it twice to any r -form results in zero. In 3D, applying the exterior derivative to various r -forms for different r corresponds to the familiar vector calculus operations of grad, div, and curl. Another important property to note is that because it operates on differential forms, it obeys a modified Leibnitz rule for wedge products. Given an r -form $v_{a_1 \dots a_r}$ and s -form $w_{b_1 \dots b_s}$, the Leibnitz rule is

$$d_c(v_{a_1 \dots a_r} \wedge w_{b_1 \dots b_s}) = (d_c v_{a_1 \dots a_r}) \wedge w_{b_1 \dots b_s} + (-1)^r v_{a_1 \dots a_r} \wedge (d_c w_{b_1 \dots b_s}). \quad (4.31)$$

Lie derivatives

Given a vector field v^a defined over the manifold, it is also possible to define a Lie derivative \mathcal{L}_v mapping a (r, s) type tensor to a (r, s) type tensor. Intuitively, the Lie derivative captures the change of a tensor along the flow of the vector field v^a . Since we will not make much use of it, we will skip the exposition for how our intuitive notion produces the form below and simply define its action on a tensor $T^{a_1 \dots a_r}_{b_1 \dots b_s}$ as follows:

$$\mathcal{L}_v T^{a_1 \dots a_r}_{b_1 \dots b_s} := v^c \partial_c T^{a_1 \dots a_r}_{b_1 \dots b_s} - \sum_{i=1}^r T^{a_1 \dots c \dots a_r}_{b_1 \dots b_s} \partial_c v^{a_i} + \sum_{j=1}^s T^{a_1 \dots a_r}_{b_1 \dots c \dots b_s} \partial_{b_j} v^c. \quad (4.32)$$

As with the exterior derivative, even though the above definition for Lie derivative uses partial derivatives, the quantity turns out to be independent of the coordinate chart that is used.

Covariant derivatives

The covariant derivative, ∇_a , formalizes the main properties of a “geometric” derivative. It acts on (r, s) type tensors to produce $(r, s + 1)$ type tensors (suggested by the Latin subscript) and satisfies the five properties listed below (with $\mu, \sigma \in \mathbb{R}$, v^a a vector, and f a function):

- (1) Obeys linearity:

$$\nabla_c(\mu A^{a_1 \dots a_r}_{b_1 \dots b_s} + \sigma B^{a_1 \dots a_r}_{b_1 \dots b_s}) = \mu \nabla_c A^{a_1 \dots a_r}_{b_1 \dots b_s} + \sigma \nabla_c B^{a_1 \dots a_r}_{b_1 \dots b_s}. \quad (4.33)$$

- (2) Obeys the Leibnitz rule:

$$\begin{aligned} \nabla_e(A^{a_1 \dots a_p}_{b_1 \dots b_q} B^{c_1 \dots c_r}_{d_1 \dots d_s}) &= (\nabla_e A^{a_1 \dots a_p}_{b_1 \dots b_q}) B^{c_1 \dots c_r}_{d_1 \dots d_s} \\ &+ A^{a_1 \dots a_p}_{b_1 \dots b_q} (\nabla_e B^{c_1 \dots c_r}_{d_1 \dots d_s}). \end{aligned} \quad (4.34)$$

- (3) Respects contraction: $\nabla_d(T^{a_1 \dots c \dots a_r}_{b_1 \dots c \dots b_s}) = \nabla_d T^{a_1 \dots c \dots a_r}_{b_1 \dots c \dots b_s}$.

- (4) Respects directional derivatives: $v^a \nabla_a f = v[f]$.

- (5) Is torsion-free: $\nabla_a \nabla_b f = \nabla_b \nabla_a f$. (Note that f is a *function*.)

The components of a tensor $\nabla_c T^{a_1 \dots a_r}_{b_1 \dots b_s}$ are often denoted by $\nabla_\lambda T^{\sigma_1 \dots \sigma_r}_{\mu_1 \dots \mu_s}$.

This is not to say we are computing the covariant derivative of the component functions of the tensor $T^{a_1 \dots a_r}_{b_1 \dots b_s}$ —we are not! It is perhaps less ambiguous to write $(\nabla_\lambda T)^{\sigma_1 \dots \sigma_r}_{\mu_1 \dots \mu_s}$, but this is not so conventional. We will adhere to the usual convention of writing the former, and take it as understood that we really mean the latter.

On a manifold, there exist many covariant derivatives satisfying the listed properties. In fact, for any coordinate chart, the associated partial derivative will satisfy the properties above; so there is a covariant derivative that computes the partial derivative associated with any coordinate chart. Often, however, ∇_a is fully specified by geometric structures (like a metric, as we shall see), so we can define and think of it in more geometric, or coordinate-free, ways.

Given the variety of covariant derivatives that exist, it then becomes interesting to consider the extent to which two covariant derivatives can differ. Perhaps surprisingly, given two different covariant derivatives (∇_a and ∇'_a), and a covector w_b , the quantity $(\nabla_a w_b - \nabla'_a w_b)$ at a point p depends only on the value of w_b at that point p . The quantity does not depend on nearby w_b even infinitesimally close to p . The quantity is also clearly (multi-)linear, so it must be a tensor. This observation allows one to write the action of any covariant derivative on a tensor as

$$\begin{aligned} \nabla_\mu T^{\sigma_1 \dots \sigma_r}_{\rho_1 \dots \rho_s} &= e_\mu [T^{\sigma_1 \dots \sigma_r}_{\rho_1 \dots \rho_s}] \\ &+ \sum_{i=1}^r A_\mu^{\sigma_i} T^{\sigma_1 \dots \lambda \dots \sigma_r}_{\rho_1 \dots \rho_s} \\ &- \sum_{j=1}^s A_\mu^\lambda T^{\sigma_1 \dots \sigma_r}_{\rho_1 \dots \lambda \dots \rho_s} \end{aligned} \quad (4.35)$$

for some tensor $A_a^b{}_c$ that depends on the choice of basis. In some texts it is common to say that $A_a^b{}_c$ is not a tensor because when one changes basis, the components do not transform according to the tensor transformation law. However, we adopt an alternate interpretation used by different texts. We say that $A_a^b{}_c$ is a tensor, but that when one changes basis, there is a completely different tensor $A'_a^b{}_c$ that must be used to compute covariant derivatives

(since the $A_a{}^b{}_c$ in Equation (4.35) depends on the choice of basis). The components are not related by the tensor transformation law because we're comparing two completely different tensors.

We have also written down a more general form than the one that appears in some introductory texts. This form is valid for noncoordinate bases as well as coordinate ones. If we choose a coordinate basis, then the action of e_μ on the coordinates of the tensor is simply given by the partial derivative ∂_μ and the tensor $A_a{}^b{}_c$ acquires a special name, $\Gamma_a{}^b{}_c$, for the Christoffel symbol. If we use an orthonormal frame (which requires the existence of a metric), we write $A_a{}^b{}_c$ as $\omega_a{}^b{}_c$ and refer to it as a set of (spin) connection coefficients. These separate symbols remind us what symmetries we can expect of the tensor. For more details on this general form, see Appendix C.1 on noncoordinate bases.

If our manifold furthermore has metric structure, we can use the metric to single out a unique covariant derivative (also referred to as the Levi-Civita connection). Intuitively, the constraint is that if we drag vectors along the manifold such that they do not exhibit any intrinsic change (technically, these vectors undergo parallel transport), then their inner products should remain the same. Mathematically, this is equivalent to requiring that

$$\nabla_c g_{ab} = 0_{cab}. \quad (4.36)$$

4.2 Intrinsic and Extrinsic Properties of Surfaces

We will now specialize our discussion to hypersurfaces embedded in an ambient manifold. Particularly, we will be interested in the classical case of surfaces in \mathbb{R}^3 . In this setting, the ambient space \mathbb{R}^3 comes with the natural Euclidean metric, denoted as g_{ab} , and therefore also possesses a unique metric-compatible covariant derivative, ∇_a . We now review two operators on vectors tangent to a surface, called fundamental forms, that together describe the surface.

4.2.1 First Fundamental Form

Classically, the first fundamental form, which operates only on vectors v^a tangent to the surface, is defined as

$$I(v^a) := g_{ab}v^av^b, \quad (4.37)$$

and returns the squared length of the vector.

The term “first fundamental form” now also commonly refers to the induced metric on the surface (which operates on two vectors, not one). If we let \hat{n}^a denote the unit normal to the surface, we can express the first fundamental form as the projection tensor:

$$P_{ab} := g_{ab} - \hat{n}_a\hat{n}_b. \quad (4.38)$$

Note that contracting the projection tensor with any two vectors (even vectors not tangent

to the surface) will compute the dot product for the surface-tangential components. We call this tensor the projection tensor because $P_a{}^b := g^{bc}P_{ac}$ takes a vector and orthogonally projects it onto the tangent plane (observe that P_{ac} is symmetric).

4.2.2 Second Fundamental Form

To introduce the second fundamental form, one classically first defines a couple related operators. The Weingarten operator is defined to be $\nabla_a \hat{n}^b$ and measures the change in normal as one moves in a direction tangent to the surface. The Weingarten operator may also be called the differential of the Gauss Map (when \hat{n}^a is viewed as a mapping from points on the surface to points on the sphere, it is called the Gauss Map). For reasons explained later, the Weingarten operator is negated and called the Shape operator: $S_a{}^b := -\nabla_a \hat{n}^b$. The second fundamental form is then defined to act on a vector v^a tangent to the surface by

$$\mathbb{II}(v^a) := g_{ab} S_c{}^a v^c v^b. \quad (4.39)$$

For unit length vectors v^a tangent to the surface, $\mathbb{II}(v^a)$ computes the *normal curvature* of the surface in the direction v^a . Unlike the case of curves in which curvature is always positive, the normal curvature is signed. The sign on the shape operator is chosen such that bending of the surface in the direction of the unit surface normal is positive, and bending in the opposite direction is negative. Some papers adopt the opposite convention, so care must be taken to translate between conventions.

Although the definition of normal curvature depends on information extrinsic to the surface (i.e., the normal), the product of minimum and maximum normal curvatures yields the *Gaussian curvature* of the surface, which is independent of the way the surface is embedded in \mathbb{R}^3 . For a more detailed discussion refer to Appendix C.3.

As with the first fundamental form, it is common nowadays to refer to the tensor being contracted against the two copies of v^a as the second fundamental form as well. So one might also define the following tensor and call it the second fundamental form:

$$\mathbb{I}_{ab} := g_{cb} S_a^c. \quad (4.40)$$

More generally, we can give a definition of the second fundamental form that relates it to the first fundamental form and is valid for families of hypersurfaces. Suppose we are given a family of hypersurfaces with a unit normal, \hat{n}^a , defined on each. We can then define the following as the second fundamental form:

$$\mathbb{I}_{ab} := -\frac{1}{2} \mathcal{L}_{\hat{n}} P_{ab} \quad (4.41)$$

$$= -\nabla_a \hat{n}_b + \hat{n}_a a_b \quad \text{where } a_b := \hat{n}^c \nabla_c \hat{n}_b. \quad (4.42)$$

Perhaps surprisingly, when this definition is applied to just a single surface, it does not matter how \hat{n}^a is extended away from the surface. Note also that higher derivatives of the normal also do not depend on the extension of the normal away from the surface as long as the derivatives are taken in directions tangent to the surface (and if we only have one surface, we won't have reason to do otherwise). We therefore do not need to be given a

normal field defined in a tubular neighborhood of our surface. Normal fields defined only on the surface are sufficient. In practice, it may be convenient to write down tensorial quantities involving derivatives of the normal before contracting against vectors tangent to the surface. In this case we do technically need a tubular neighborhood, but any arbitrary extension will do, as long as we remember to only contract in surface tangent directions afterward or use these tensors in ways that only relate to the given surface (so do not depend on the extension).

Since the second fundamental form (and higher order derivatives of normal) do not depend on how we extend the vector field away from the surface, for computations we might as well extend vector fields by parallel transport. This causes the expressions above to simplify:

$$\mathbb{I}_{ab} := -\frac{1}{2}\mathcal{L}_{\hat{n}}\mathcal{P}_{ab} \quad (4.43)$$

$$= -\frac{1}{2}\mathcal{L}_{\hat{n}}g_{ab} \quad (4.44)$$

$$= -\nabla_a\hat{n}_b. \quad (4.45)$$

From the definition in Equation (4.41), we can see that \mathbb{I}_{ab} is manifestly symmetric. This is less obvious from the more classical definition in Equation (4.40).

4.2.3 Tensors that live on the surface itself

Although the first and second fundamental forms introduced in the previous sections were defined using the ambient space, the tensors really live on the surface itself. The first

fundamental form is a metric for vectors tangent to the surface. The second fundamental form maps vectors tangent to the surface to vectors tangent to the surface. We can formalize these ideas using the notions of push-forward and pull-back.

Given a mapping f between manifolds, we have an induced mapping $(df)_{\bar{a}}^b$ between the tangent spaces of the domain and range. In coordinates, this is represented by the Jacobian of the mapping. We will adorn objects associated with the domain with a bar (over-line) and leave objects associated with the range unadorned. This convention carries over to the indices associated with the objects. So given f , we can “push” a vector $\bar{v}^{\bar{a}}$ in the domain’s tangent space to a vector v^b in the range’s tangent space via the function’s differential $(df)_{\bar{a}}^b$:

$$v^b := (df)_{\bar{a}}^b \bar{v}^{\bar{a}}. \quad (4.46)$$

Likewise, given a covector w_b in the cotangent space of the range, we can “pull” it back to a covector $\bar{w}_{\bar{a}}$ in the cotangent space of the domain:

$$\bar{w}_{\bar{a}} := (df)_{\bar{a}}^b w_b. \quad (4.47)$$

In general, we can always push-forward $(r, 0)$ type tensors and pull-back $(0, s)$ type tensors. Specifically, if f is an embedding of a surface into \mathbb{R}^3 , we can pull back the metric from \mathbb{R}^3 to get $\bar{g}_{\bar{a}\bar{b}}$. We can also pull back the second fundamental form to get $\bar{\Pi}_{\bar{a}\bar{b}}$. There is also a unique covariant derivative on the surface compatible with the pulled back metric and we denote it as $\bar{\nabla}_{\bar{a}}$. This covariant derivative can be computed in \mathbb{R}^3 by first applying the 3D covariant derivative, and then projecting the resulting tensor onto the tangent and cotangent

spaces of the surface.

In the more classical notation, given a parameterization of the surface, $X(u, v)$, into \mathbb{R}^3 , the components of the surface metric are computed by

$$\mathbf{I} = \begin{bmatrix} X_u \cdot X_u & X_u \cdot X_v \\ X_v \cdot X_u & X_v \cdot X_v \end{bmatrix}. \quad (4.48)$$

where the subscripts u and v denote partial differentiation with respect to the marked variable. The components of the second fundamental form are computed by

$$\mathbf{II} = \begin{bmatrix} \hat{n} \cdot X_{uu} & \hat{n} \cdot X_{uv} \\ \hat{n} \cdot X_{vu} & \hat{n} \cdot X_{vv} \end{bmatrix} \quad (4.49)$$

$$= \begin{bmatrix} -\hat{n}_u \cdot X_u & -\hat{n}_u \cdot X_v \\ -\hat{n}_v \cdot X_u & -\hat{n}_v \cdot X_v \end{bmatrix}. \quad (4.50)$$

These relations are presented only for completeness; we will not make use of this notation.

4.2.4 Higher Order Derivatives

When studying surfaces, we will also have reason to look at higher order derivatives. Here we specialize our discussion to higher order derivatives of normal using an orthonormal frame. Note that this choice of frame is an example of a noncoordinate basis. In an orthonormal frame, the metric always has components matching the identity matrix. However, it is impossible to choose a chart in which the metric over any local region is constant unless the manifold itself has no Gaussian curvature in that region (in a coordinate basis,

the metric's components and their derivatives reflect the manifold's intrinsic curvature). Therefore, this choice of orthonormal frame cannot come from a coordinate chart. Instead, when using an orthonormal frame adapted to a surface (i.e., two basis vectors lie tangent to the surface), information about the surface's intrinsic curvature is captured by how the frame rotates as one transitions from point to point on the manifold.

The rotation of the orthonormal frame field is captured by the (spin) connection coefficients, $\omega_a^b{}_c$. These are the same connection coefficients that appear when computing the covariant derivative in the orthonormal basis (see Appendix C.1). We define the connection coefficients as follows:

$$\omega_a^\mu{}_\nu := (\vartheta^\mu)_b \nabla_a (e_\nu)^b. \quad (4.51)$$

These connection coefficients are not allowed to be arbitrary. For example, one consequence of orthonormality of the frame is given by the following:

$$\omega_a^\mu{}_\nu = -\omega_a^\nu{}_\mu, \quad \forall \mu, \nu. \quad (4.52)$$

In order for them to describe a valid surface, they must also obey the *structure equations*. For details, refer to Appendix C.2. Values for particular choices of orthonormal frame are given in Sections 5.2.1 and 5.2.2.

Once the connection coefficient values are known, the rule for applying covariant derivatives can be used to obtain higher order derivatives of normals (or other tensors). For

example, the second derivative of normal can be computed by the following expression:

$$\nabla_\mu \nabla_\nu \hat{n}^\lambda = \nabla_\mu (\nabla_\nu \hat{n}^\lambda) \quad (4.53)$$

$$= e_\mu [\omega_\nu^\lambda] + \omega_\mu^\lambda \omega_\nu^1 \omega_1^3 + \omega_\mu^\lambda \omega_\nu^2 \omega_2^3 - \omega_\mu^1 \omega_\nu \omega_1^\lambda - \omega_\mu^2 \omega_\nu \omega_2^\lambda. \quad (4.54)$$

We emphasize that Equation (4.54) is not valid for *any* choice of orthonormal frames. It is valid only for orthonormal frames adapted to the surface (i.e., the third basis vector corresponds exactly to the surface normal). Likewise, higher order derivatives can be computed as well:

$$\nabla_\mu \nabla_\nu \nabla_\rho \hat{n}^\lambda = \nabla_\mu (\nabla_\nu \nabla_\rho \hat{n}^\lambda) \quad (4.55)$$

$$= e_\mu [\nabla_\nu \nabla_\rho \hat{n}^\lambda] + \omega_\mu^\lambda \omega_\sigma \nabla_\nu \nabla_\rho \hat{n}^\sigma - \omega_\mu^\sigma \omega_\nu \nabla_\sigma \nabla_\rho \hat{n}^\lambda - \omega_\mu^\sigma \omega_\rho \nabla_\nu \nabla_\sigma \hat{n}^\lambda. \quad (4.56)$$

4.3 The Metric Equivalence Problem

In differential geometry, two mathematical structures are considered equivalent if they can be related by some smooth reparameterization. Given that reparameterizations are allowed to be arbitrarily complicated, determining equivalence is not always an easy problem. In the case of metrics, this is known as the *metric equivalence problem*. For Riemannian metrics, the problem is stated as follows: given two metrics g_{ab} and g'_{ab} , when does there exist a smooth map f (whose inverse is also smooth) such that one metric is the pullback

of the other? I.e.,

$$g_{ab} = (df)_a^c (df)_b^d g'_{cd}. \quad (4.57)$$

For Riemannian metrics, it turns out the question can be answered in terms of differential invariants derived from the Riemann curvature tensor. The Riemann curvature tensor itself is computable from the metric. One consequence is that if two manifolds possess different Riemannian curvatures, the difference provides an obstruction to their equivalence. This shows, for example, that a Euclidean space is not equivalent to any space with some amount of nonzero intrinsic curvature.

Chapter 5

Shapes from Curves

In this chapter, we consider the problem of inferring shape from curves in images. As shown in Figure 1.4, curves alone are often enough to engender shape understanding. However, not all types of curves are equally evocative. We single out suggestive contours and illumination valleys as particularly interesting because although one is defined in terms of three-dimensional geometry and the other in terms of image features, the two produce strikingly similar (and effective) results. This suggests that the two types of curves capture and convey similar pieces of geometric information. To explore this connection, we develop some general techniques for recasting questions about the image as questions about the surface. The methods allow us to prove some initial relations between curves and three-dimensional geometry. However, fully characterizing the relationships between curves and shape is difficult and the current analysis leaves many basic and important questions still unanswered. We hope that the groundwork we lay will prove useful for future analyses.

5.1 Introduction

We seek to relate curves in an image to three-dimensional geometric information. As shown in Figure 1.4, curves alone are often enough to convey a sense of shape. This naturally raises the following question: What curves do humans actually latch onto, and what precisely is their geometric information content?

Since human vision can only make use of retinal inputs, curve understanding is plausibly constructed using only information that is stably detected in images. It is therefore worth noting that suggestive contours—the curves displayed in Figures 1.4a and 5.1—tend to line up surprisingly well with valleys in the image’s illumination function (when viewing a diffuse object lit by a light co-located with the view camera) [15]. This connection between illumination valleys and suggestive contours is interesting because illumination valleys are defined (and detectable) in the image, whereas suggestive contours are defined in terms of surface properties. The empirical match occurs frequently and persists under many different views, hinting that the correspondence is more than mere coincidence. It is perhaps the case that suggestive contours are suggestive precisely because of this relation to stably detectable image properties.

In this chapter, we explore connections between suggestive contours, illumination valleys, and shading. An initial attempt at analyzing the connection might start by representing the surface as a heightfield over the image plane. This allows one to write down geometric relations in terms of an unknown height at each point. Unfortunately such a representation is not conducive to simple analysis: it tends to lead to many terms when computing geometric quantities of interest. For example, even in factored form, the expres-

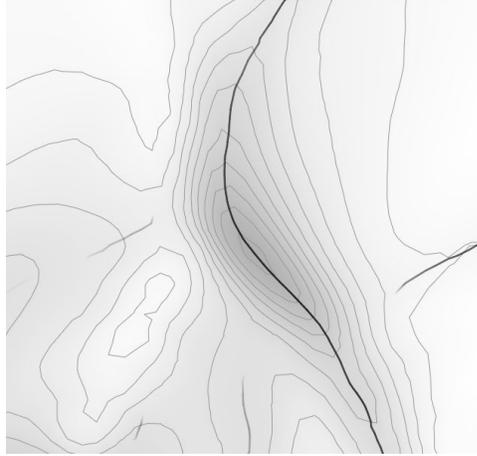


Figure 5.1: Suggestive contours (thick black line) tend to line up with illumination valleys in the image. We also draw the iso-luminance contours to help visualize the shading.

sion for Gaussian curvature in terms of height (z) is not particularly simple:

$$\frac{1}{(1 + z_x^2 + z_y^2)^3} \left\{ (z_{xx}(1 + z_y^2) - z_x z_y z_{xy})(z_{yy}(1 + z_x^2) - z_x z_y z_{xy}) \right. \\ \left. - (z_{xy}(1 + z_y^2) - z_x z_y z_{yy})(z_{xy}(1 + z_x^2) - z_x z_y z_{xx}) \right\}. \quad (5.1)$$

Faced with an expression that relates multiple geometric quantities, it would then be difficult to find the right rearrangement and factorization of terms to identify interesting geometric quantities.

Instead, by adapting the bases in which we represent information to the geometric configuration of the scene, we can simplify many expressions while still capturing important geometric relations. The differential geometry of Chapter 4 becomes useful because at each point on the surface or image, we will choose different bases in which to carry out computations—we do not use one global basis in which to do all our work. In particular, this means that when taking derivatives, we cannot simply use partial or directional derivatives.

Measuring change in such a way would capture not only geometric changes in quantities of interest, but also arbitrary changes due to the rotations of our bases. We employ a covariant derivative instead to compute a “geometric” derivative that discounts the effect of changing bases.

We start our discussion by introducing terminology and defining curves of various types. In addition to the suggestive contour and the Saint-Venant valley, we define a couple additional curves as well; these are provided both for completeness and for relating suggestive contours and Saint-Venant curves to other studied and informative sources of information. We then analyze the relationship between three-dimensional geometry and image projections via a convenient choice of basis. Finally, we discuss the behavior of suggestive contours and Saint-Venant curves at critical points of illumination and prove some more limited results about these curves away from such points.

5.2 Curve Definitions

Curves on surfaces can be classified according to a number of divisions. There are those curves that are defined in terms of the geometric properties of the surface and its embedding (i.e., intrinsic and extrinsic curvatures) only. Then there are those curves that require additional functions to be defined over the surface. These functions provide information that relates the surface geometry to other geometric objects of interest. As examples of these, we focus on view dependence (auxiliary geometric object is a vantage point) and illumination dependence (auxiliary geometric object is an external light source).

5.2.1 Surface-only Curves

We first present a special orthonormal frame, the principal frame, and give some relations that are useful for performing computations in such a frame (these relations are mainly just for reference). We then define some of the more common curves that depend only on the surface and its embedding.

Principal frame

To deal with curves that depend only on the surface's geometry and its embedding into \mathbb{R}^3 , it is handy to use the *principal frame*. In this orthonormal frame, two basis vectors are chosen to align with the two principal directions on the surface (directions of maximum and minimum normal curvature) and the third basis vector aligns with the unit (outward-pointing) normal.

Let κ_1 and κ_2 denote the two principal curvatures (largest and smallest normal curvatures). Then the components for the first derivative of normal (i.e., the Weingarten operator) in the principal frame take a simple form:

$$\nabla_\mu \hat{n}^\nu = \begin{bmatrix} \omega_1^1{}_3 & \omega_2^1{}_3 & \omega_3^1{}_3 \\ \omega_1^2{}_3 & \omega_2^2{}_3 & \omega_3^2{}_3 \\ \omega_1^3{}_3 & \omega_2^3{}_3 & \omega_3^3{}_3 \end{bmatrix} = \begin{bmatrix} -\kappa_1 & 0 & 0 \\ 0 & -\kappa_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (5.2)$$

Recall that the covariant derivative here, ∇_a , is the covariant derivative associated with \mathbb{R}^3 . This is necessitated by the fact that the normal vector \hat{n}^b is not tangent to the surface.

The connection coefficients $\omega_\mu^\nu{}_\lambda$ are associated with the chosen basis (here, the principal frame). In Section 5.2.2, where we use the radial frame instead, we get correspondingly different connection coefficients.

The components for the second covariant derivative of the surface normal, computed by applying Equation (4.54), are given by the following (where the three matrices are stacked on the right to correspond to the upper index values of one, two, and three respectively):

$$\begin{bmatrix} \nabla_1 \nabla_1 \hat{n} & \nabla_1 \nabla_2 \hat{n} & \nabla_1 \nabla_3 \hat{n} \\ \nabla_2 \nabla_1 \hat{n} & \nabla_2 \nabla_2 \hat{n} & \nabla_2 \nabla_3 \hat{n} \\ \nabla_3 \nabla_1 \hat{n} & \nabla_3 \nabla_2 \hat{n} & \nabla_3 \nabla_3 \hat{n} \end{bmatrix} = \begin{bmatrix} -e_1[\kappa_1] & -e_2[\kappa_1] & -(\kappa_1)^2 \\ -e_2[\kappa_1] & -e_1[\kappa_2] & 0 \\ -(\kappa_1)^2 & 0 & 0 \end{bmatrix} \quad (5.3)$$

$$\begin{bmatrix} -e_2[\kappa_1] & -e_1[\kappa_2] & 0 \\ -e_1[\kappa_2] & -e_2[\kappa_2] & -(\kappa_2)^2 \\ 0 & -(\kappa_2)^2 & 0 \end{bmatrix} \quad (5.4)$$

$$\begin{bmatrix} -(\kappa_1)^2 & 0 & 0 \\ 0 & -(\kappa_2)^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (5.5)$$

Recall that for these tensors, we should only contract the indices notationally coming from derivatives with directions tangent to the surface or use these tensors in ways such that the overall expression does not depend on the normal's extension away from the surface.

Derivatives of normal, along with the commutativity of the covariant derivative in \mathbb{R}^3 (i.e., $\nabla_a \nabla_b = \nabla_b \nabla_a$), provide further relations between connection coefficients and

surface geometry. These facts can also be derived using the structure equations in Appendix C.2. These are included here largely for reference and can be safely skipped. To present them, we first define the commutator of two basis vectors (which is itself a vector):

$$[e_1, e_2]^a := (\omega_1^\mu{}_2 - \omega_2^\mu{}_1)(e_\mu)^a. \quad (5.6)$$

Then the structural relations are as follows:

$$(\kappa_1 - \kappa_2)[e_1, e_2][\kappa_1] = -e_2[\kappa_1]e_1[\kappa_1 + \kappa_2], \quad (5.7)$$

$$(\kappa_1 - \kappa_2)[e_1, e_2][\kappa_2] = -e_1[\kappa_2]e_2[\kappa_1 + \kappa_2], \quad (5.8)$$

$$\begin{aligned} (\kappa_1 - \kappa_2)(e_1[e_1[\kappa_2]] - e_2[e_2[\kappa_1]]) &= -2(e_1[\kappa_2])^2 - 2(e_2[\kappa_1])^2 \\ &\quad + e_1[\kappa_1]e_1[\kappa_2] + e_2[\kappa_1]e_2[\kappa_2] \\ &\quad - \kappa_1 \kappa_2 (\kappa_1 - \kappa_2)^2. \end{aligned} \quad (5.9)$$

The structure equations can also be used to derive the simple relations below:

$$e_1[\kappa_2] = (\kappa_1 - \kappa_2)\omega_2^2{}_1, \quad (5.10)$$

$$e_2[\kappa_1] = (\kappa_1 - \kappa_2)\omega_1^2{}_1. \quad (5.11)$$

Parabolic Curves

Parabolic points are those where the Gaussian curvature of the surface is zero:

$$K := \kappa_1 \kappa_2 = 0. \quad (5.12)$$

On a generic closed surface, the collection of these points form closed loops. These loops separate elliptic (sphere-like) regions of the surface from hyperbolic (saddle-like) regions of the surface. As we shall see later, parabolic lines can be quite informative, especially in relation to critical points of illumination.

Ridges and Valleys

Ridge and valley points are those where one of the principal curvatures becomes extremal in its corresponding principal direction. In the principal frame, the condition can be written as follows:

$$e_1[\kappa_1] e_2[\kappa_2] = 0. \quad (5.13)$$

Whether points are classified as ridges or valleys depends on whether the extrema is a maximum or minimum. Let $(e_\mu)^a$ be the principal direction in which the derivative of its corresponding principal curvature is zero. If $e_\mu [e_\mu[\kappa_\mu]]$ is positive, then the critical point is a valley. If $e_\mu [e_\mu[\kappa_\mu]]$ is negative, then the critical point is a ridge.

Ridges and valleys provide useful information about a surface's shape, but simply plotting them on a surface is not generally considered particularly evocative from a perceptual standpoint.

Mean Curvature

Mean curvature curves are defined to be the set of points where mean curvature is zero.

$$H := \frac{1}{2}(\kappa_1 + \kappa_2) = 0. \quad (5.14)$$

Mean curvature curves can be given intuitive meaning as follows. Imagine the surface as a fluid interface or evolving soap film; then the pressure (force per area) across the interface at any point is directly proportional to the mean curvature. Therefore, any point at which the mean curvature is zero is likewise a point where the pressure on the surface is zero. Infinitesimally, these locations are stable as the surface evolves under mean curvature flow (e.g., the soap film evolves toward an equilibrium state).

5.2.2 Environment-dependent Curves

In this section, we introduce two geometric objects external to the surface: a view point and a light source. For simplicity, we will assume the view camera captures an orthographic projection of the scene and the light source is a point at infinity (i.e., is directional). These additional geometric objects provide extra pieces of information that can be used to select preferred directions beyond the principal directions. We discuss new choices of bases and then define curves that not only depend on surface geometry but also on its relationship with respect to the other geometric objects.

Radial frame

The presence of a camera viewpoint allows us to construct a new preferred orthonormal frame involving both surface and view information. Let C be the camera position.¹ Let p be a point on the surface. Let $V^a := C - p$ be the vector pointing from the point p to the camera. Let \hat{v}^a be its normalization as a vector in \mathbb{R}^3 . The unnormalized radial vector is defined as $r^a := P_b^a V^b$ (P_b^a , defined in Equation (4.38), kills the normal component of a vector). The normalized radial direction is $\hat{r}^a := \frac{r^a}{\|r\|_g}$, where $\|\cdot\|_g$ denotes the length with respect to the metric g_{ab} . Define \hat{r}_\perp^a to be the (positive) ninety degree rotation of \hat{r}^a about the unit outward pointing normal. Then the ordered triple $(\hat{r}^a, \hat{r}_\perp^b, \hat{n}^c)$ forms a right-handed orthonormal frame that we call the *radial frame*.

To effectively carry out computations in this frame, we need to determine the connection coefficients. These are computed by taking the covariant derivative of each basis vector. We get the following for the basis vectors tangent to the surface:

$$\hat{r}^a \nabla_a \hat{r}^b = \omega_1^2 \hat{r}_\perp^b + \omega_1^3 \hat{n}^b = \tau_r \cot \alpha \hat{r}_\perp^b + \kappa_r \hat{n}^b \quad (5.15)$$

$$\hat{r}_\perp^a \nabla_a \hat{r}^b = \omega_2^2 \hat{r}_\perp^b + \omega_2^3 \hat{n}^b = \bar{\kappa}_r \cot \alpha \hat{r}_\perp^b + \tau_r \hat{n}^b \quad (5.16)$$

$$\hat{r}^a \nabla_a \hat{r}_\perp^b = \omega_1^1 \hat{r}_\perp^b + \omega_1^3 \hat{n}^b = -\tau_r \cot \alpha \hat{r}_\perp^b + \tau_r \hat{n}^b \quad (5.17)$$

$$\hat{r}_\perp^a \nabla_a \hat{r}_\perp^b = \omega_2^1 \hat{r}_\perp^b + \omega_2^3 \hat{n}^b = -\bar{\kappa}_r \cot \alpha \hat{r}_\perp^b + \bar{\kappa}_r \hat{n}^b \quad (5.18)$$

The connection coefficients $\omega_\mu^{\nu\lambda}$ here are *different* from those in Section 5.2.1 because we are using the radial frame instead of the principal frame. The angle α denotes the

¹Since we are assuming an orthographic camera, the pinhole C should be taken to be a point at infinity. For these constructions, we can instead first assume C is finite and take the limit as the point approaches a point at infinity. In this way, the normalized vectors are still well defined.

angle between the surface normal and the viewing direction. The symbol κ_r , called the *radial curvature*, is the normal curvature of the surface in the radial direction. The symbol τ_r is the geodesic torsion associated with the radial direction. If η is the angle between the radial direction and a principal direction, then the geodesic torsion is equal to $(\kappa_1 - \kappa_2) \sin \eta \cos \eta$. For principal directions, the geodesic torsion is zero. The symbol $\bar{\kappa}_r$ is the normal curvature in the \hat{r}_\perp^a direction. The derivation of these facts is presented in Appendix D.1. From the antisymmetry in the last two indices of the connection coefficients, we can also write out the components of the Weingarten operator:

$$\nabla_\mu \hat{n}^\nu = \begin{bmatrix} \omega_1^1{}_3 & \omega_2^1{}_3 & \omega_3^1{}_3 \\ \omega_1^2{}_3 & \omega_2^2{}_3 & \omega_3^2{}_3 \\ \omega_1^3{}_3 & \omega_2^3{}_3 & \omega_3^3{}_3 \end{bmatrix} = \begin{bmatrix} -\kappa_r & -\tau_r & 0 \\ -\tau_r & -\bar{\kappa}_r & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (5.19)$$

Up to antisymmetry in the last two indices, these are all the nonzero connection coefficients. Again, these connection coefficients differ from those in Section 5.2.1 since we are using the radial frame.

As with the principal frame, we can discover more relations by computing one more order derivative of normal and equate the components that should be identified because of the commutativity of the covariant derivative in \mathbb{R}^3 . As in [15], we may denote the second derivative of normal by $C_{ab}{}^d := \nabla_a \nabla_b \hat{n}^d$. Again, these details are provided for

reference and can be safely skipped.

$$C_{\mu\nu}{}^\sigma = [\nabla_\mu \nabla_\nu \hat{n}]^\sigma \quad (5.20)$$

$$= \begin{bmatrix} -e_1[\kappa_r] + 2\tau_r^2 \cot \alpha & -e_1[\tau_r] - \tau_r(\kappa_r - \bar{\kappa}_r) \cot \alpha & -\kappa_r^2 - \tau_r^2 \\ -e_2[\kappa_r] + 2\bar{\kappa}_r \tau_r \cot \alpha & -e_2[\tau_r] - \bar{\kappa}_r(\kappa_r - \bar{\kappa}_r) \cot \alpha & -2\tau_r H \\ -\kappa_r^2 - \tau_r^2 & -2\tau_r H & 0 \end{bmatrix} \quad (5.21)$$

$$\begin{bmatrix} -e_1[\tau_r] - \tau_r(\kappa_r - \bar{\kappa}_r) \cot \alpha & -e_1[\bar{\kappa}_r] - 2\tau_r^2 \cot \alpha & -2\tau_r H \\ -e_2[\tau_r] - \bar{\kappa}_r(\kappa_r - \bar{\kappa}_r) \cot \alpha & -e_2[\bar{\kappa}_r] - 2\bar{\kappa}_r \tau_r \cot \alpha & -\tau_r^2 - \bar{\kappa}_r^2 \\ -2\tau_r H & -\tau_r^2 - \bar{\kappa}_r^2 & 0 \end{bmatrix} \quad (5.22)$$

$$\begin{bmatrix} -\kappa_r^2 - \tau_r^2 & -2\tau_r H & 0 \\ -2\tau_r H & -\bar{\kappa}_r^2 - \tau_r^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (5.23)$$

The nonzero components of the tensor and their symmetries are listed below.

$$C_{11}{}^1 \quad (5.24)$$

$$C_{12}{}^1 = C_{11}{}^2 = C_{21}{}^1 \quad (5.25)$$

$$C_{21}{}^2 = C_{12}{}^2 = C_{22}{}^1 \quad (5.26)$$

$$C_{22}{}^2 \quad (5.27)$$

$$C_{11}{}^3 = C_{31}{}^1 = C_{13}{}^1 \quad (5.28)$$

$$C_{22}{}^3 = C_{32}{}^2 = C_{23}{}^2 \quad (5.29)$$

$$C_{21}{}^3 = C_{12}{}^3 = C_{32}{}^1 = C_{23}{}^1 = C_{31}{}^2 = C_{13}{}^2 \quad (5.30)$$

The symmetries give us the following relations:

$$-e_1[\bar{\kappa}_r] + K \cot \alpha = -e_2[\tau_r] + (\bar{\kappa}_r^2 + \tau_r^2) \cot \alpha \quad (5.31)$$

$$e_1[\tau_r] + 2H\tau_r \cot \alpha = e_2[\kappa_r] \quad (5.32)$$

Imaging Plane

The existence of a view camera also gives us an imaging plane, or image screen, on which we see an orthographic projection of our surface. This image screen comes with a natural choice of basis, the Cartesian coordinate basis vectors for \mathbb{R}^2 , which we will denote by $(\frac{\partial}{\partial \check{x}^{\bar{\mu}}})^{\bar{a}}$. We will denote its dual basis as $(d\check{x}^{\bar{\mu}})_{\bar{a}}$. The screen also comes with a natural Euclidean metric $\check{g}_{\bar{a}\bar{b}} := \bar{\delta}_{\bar{\mu}\bar{\nu}}(d\check{x}^{\bar{\mu}})_{\bar{a}}(d\check{x}^{\bar{\nu}})_{\bar{b}}$. We will use the bar (over-line) over variable names and indices to denote that they live on the screen. The breve ($\check{}$) will be used over variable names for objects that have been normalized to unit length using the screen metric or are otherwise intimately tied to the screen metric.

The image screen can also be thought of as a parameterization domain for the visible parts of the surface (minus the silhouette points). This is what is done when we think of the surface as a heightfield over the imaging plane. As discussed in Section 4.2.3, we can then “pull” many of the tensors we think of as living intrinsically on the surface back to the screen (i.e., the parameter domain) via the surface parameterization. For example, we can pull back the metric on the surface to get a corresponding tensor $\bar{g}_{\bar{a}\bar{b}}$ living on the screen. Likewise, for vectors $\bar{v}^{\bar{a}}$ on the screen, we can push them forward to vectors living in the tangent space of the surface (which is isomorphic to \mathbb{R}^2). The inclusion map that

takes a surface and identifies it with its embedding in \mathbb{R}^3 can then be used to push the vector forward again to a vector living in \mathbb{R}^3 (where its component in the normal direction is just zero). We denote this vector as v^a . These notational conventions, then, are largely just different ways of thinking about the same geometric objects.

One upshot is that on the screen we have two metrics—one being the metric associated with the surface, and the other being the metric associated with the screen itself (treating the screen as a Euclidean space). The former is more informative in terms of the geometry, but the latter is the only one we have direct access to (and would use to make inferences necessary for learning the former). Given these two metrics (which we call the “surface metric” and “screen metric”), we also have two corresponding metric compatible covariant derivatives living on the screen. Let $\bar{\nabla}_{\bar{a}}$ denote the covariant derivative compatible with the surface metric. Let $\check{\nabla}_{\check{a}}$ denote the covariant derivative compatible with the screen metric.

Illumination Function

If we are given a light source, we can also define an illumination function on the surface. Analogous to the view point case, let E be the light position, and let $L^a := E - p$ be the vector pointing from the point p to the light.² Denote the \mathbb{R}^3 -normalized light vector as $\hat{\ell}^a$. We then define the illumination function as $\cos \theta := \hat{n}^a \hat{\ell}_a$.

For our analysis, we will be concerned with the “miner’s headlamp” case, where the light position is the same as the camera position. In this case, $\theta = \alpha$, and $\phi = 0$ (see

²As with the view point case, we can treat E as a finite point and then take the limit as it approaches a point at infinity.

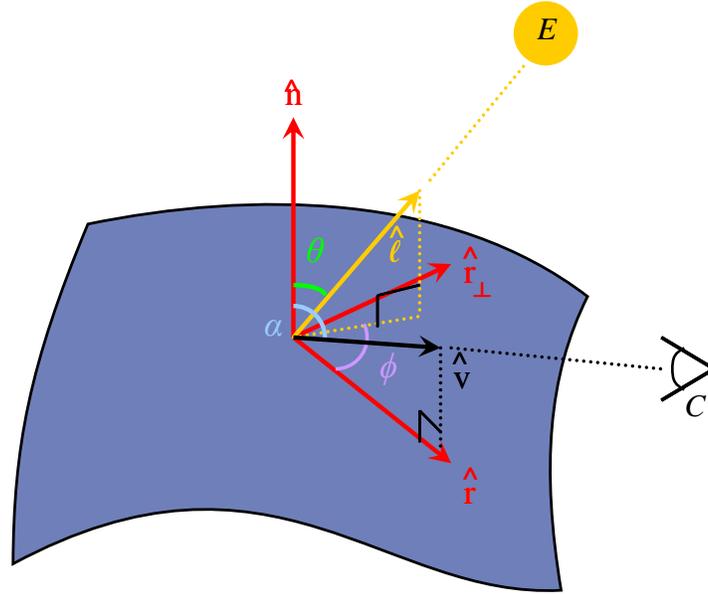


Figure 5.2: Relations between surface, view direction, and light direction. For miner's headlamp, the light source is co-located with the camera pinhole, so $\theta = \alpha$ and $\phi = 0$.

Figure 5.2). The angle ϕ is the angle between the radial direction and the light vector's projection onto the tangent plane.

Given the illumination function, we can define various objects that are naturally detectable in the image. Let us define the illumination gradient vector by

$$\bar{\xi}^{\bar{a}} := g^{\bar{a}\bar{b}} \bar{\nabla}_{\bar{b}} \cos \theta \quad (5.33)$$

$$= g^{\bar{a}\bar{b}} \bar{\nabla}_{\bar{b}} \cos \theta. \quad (5.34)$$

The equality of the second line follows from the fact that the various types of derivatives are all required to act the same on scalar functions.

The isophote vector $\bar{\xi}_{\perp}^{\bar{a}}$ is simply the 90 degree right-handed rotation of $\bar{\xi}^{\bar{a}}$ in the image plane. Note that \perp here does not imply perpendicular on the surface (using the

surface metric), but perpendicular on screen (using the screen metric). When normalized with respect to the surface metric, we will denote these vectors as $\hat{\xi}^{\bar{a}}$ and $\hat{\xi}_{\perp}^{\bar{a}}$ respectively. Likewise, when normalized with respect to the screen metric, we will denote these vectors as $\check{\xi}^{\bar{a}}$ and $\check{\xi}_{\perp}^{\bar{a}}$. We will avoid writing these vectors with a lower latin index to avoid the ambiguity of whether it was lowered using the screen metric or the surface metric.

We define the illumination hessian as follows:

$$\check{H}_{\bar{a}}^{\bar{b}} := \check{g}^{\bar{b}\bar{c}} \check{\nabla}_{\bar{c}} \check{\nabla}_{\bar{a}} \cos \theta. \quad (5.35)$$

The hessian is often defined in terms of partial derivatives. The above is a more geometric way of defining the same object without explicit reference to a specifically chosen coordinate system. This perspective is useful when we want to compute the components of the same geometric object, but in a different basis. We have also used the screen metric to raise one of the indices so that we have a type $(1, 1)$ object, and can then define eigenvalues and eigenvectors. This is the typical interpretation of the illumination “hessian matrix.”

Suggestive Contours

Suggestive contours are a view dependent set of curves [15]. They are formally defined as the set of points where radial curvature vanishes:

$$\kappa_r := \Pi_{ab} \hat{r}^a \hat{r}^b = 0. \quad (5.36)$$

A subset of the zero set is culled to leave only the suggestive contours corresponding to “valleys.” Curves are drawn only if the following condition is upheld:

$$\hat{r}[\kappa_r] < \varepsilon, \quad (5.37)$$

where ε is some small threshold value. As illustrated in Figure 1.4, they are quite effective at conveying some shapes.

Apparent Ridges

Apparent ridges are another set of view dependent curves that are effective at conveying some shapes [32]. They are based on the intuition that the definition of ridges should depend on the viewer’s judgment of surface curvature. Let κ'_1 be the largest eigenvalue (in magnitude) of $g^{ab}\bar{\Pi}_{bc}$ and let \bar{v}^{1a} be its corresponding eigenvector. Then the apparent ridges are the points where:

$$\bar{v}^{1a}[\kappa'_1] = 0 \quad \text{and} \quad \kappa'_1 \cdot \bar{v}^{1a}[\bar{v}^{1a}[\kappa'_1]] < 0 \quad (5.38)$$

The sign condition tests that the magnitude of the curvature is maximized at the critical point. Curves can also be classified as ridge or valley by the sign of the curvature at the critical point. If $\kappa'_1 < 0$, then the point is a ridge; if $\kappa'_1 > 0$, then the point is a valley.

Saint-Venant

Saint-Venant curves are the set of points where the illumination gradient is zero or is an eigenvector of the illumination hessian matrix:

$$\check{g}_{\bar{c}\bar{b}} \check{H}_{\bar{a}}^{\bar{b}} \bar{\xi}^{\bar{a}} \bar{\xi}_{\perp}^{\bar{c}} = 0. \quad (5.39)$$

Alternatively, the Saint-Venant curves can also be defined as the points where the gradient magnitude reaches a critical point in the isophote direction:

$$\bar{\xi}_{\perp} [\|\bar{\xi}\|_{\check{g}}^2] = 0. \quad (5.40)$$

The equivalence can be seen as follows:

$$\bar{\xi}_{\perp} [\|\bar{\xi}\|_{\check{g}}^2] = \bar{\xi}_{\perp}^{\bar{a}} \check{\nabla}_{\bar{a}} (\check{g}_{\bar{b}\bar{c}} \bar{\xi}^{\bar{b}} \bar{\xi}^{\bar{c}}) \quad (5.41)$$

$$= 2 \check{g}_{\bar{b}\bar{c}} \bar{\xi}^{\bar{b}} \bar{\xi}_{\perp}^{\bar{a}} \check{\nabla}_{\bar{a}} (\bar{\xi}^{\bar{c}}) \quad (5.42)$$

$$= 2 \check{g}_{\bar{b}\bar{c}} \bar{\xi}^{\bar{b}} \bar{\xi}_{\perp}^{\bar{a}} \check{H}_{\bar{a}}^{\bar{c}}. \quad (5.43)$$

Using the symmetry of $\check{H}_{\bar{a}}^{\bar{c}}$ and the fact that the extra factor of 2 does not affect the zero set, we see that the expression in (5.43) matches the left-hand side of Equation (5.39).

On a Saint-Venant curve, the gradient and isophote directions are eigenvectors of the illumination hessian. Let λ and λ_{\perp} be the eigenvalues corresponding to $\bar{\xi}^{\bar{a}}$ and $\bar{\xi}_{\perp}^{\bar{a}}$ respectively. The illumination valleys and ridges occur when $|\lambda| < |\lambda_{\perp}|$. We can further distinguish valley from ridge by checking the sign of λ_{\perp} . If $\lambda_{\perp} > 0$, then we classify the

point as a Saint-Venant valley; otherwise, we label the point as a Saint-Venant ridge.

Canny

Canny edges are defined as points where the illumination gradient reaches a critical point in the gradient direction:

$$\bar{\xi} [\|\bar{\xi}\|_{\bar{g}}] = 0. \quad (5.44)$$

5.3 Relations Between Image and Surface Curves

5.3.1 Image Plane and Orthographic Projection

For analysis, we want to relate curves in the image (information that is available) to properties of the surface (information we want to learn).

The radial frame provides a convenient basis for relating quantities on the surface to those in the camera's image screen since the two surface-tangential basis vectors, which are orthogonal on the surface, stay orthogonal when projected onto the screen (see Figure 5.3). Define the two screen vectors $\check{r}^{\bar{a}}$ and $\check{r}_{\perp}^{\bar{a}}$ as the screen-orthographic projections of $\hat{r}^{\bar{a}}$ and $\hat{r}_{\perp}^{\bar{a}}$ that are normalized according to the screen metric. As pointed out in Section 5.2.2, we can just as well think of these vectors—as well as other vectors defined on the screen—as vectors living in \mathbb{R}^3 , so we will write \check{r}^a and \check{r}_{\perp}^a to avoid the notational clutter of having bars over every variable name and index.

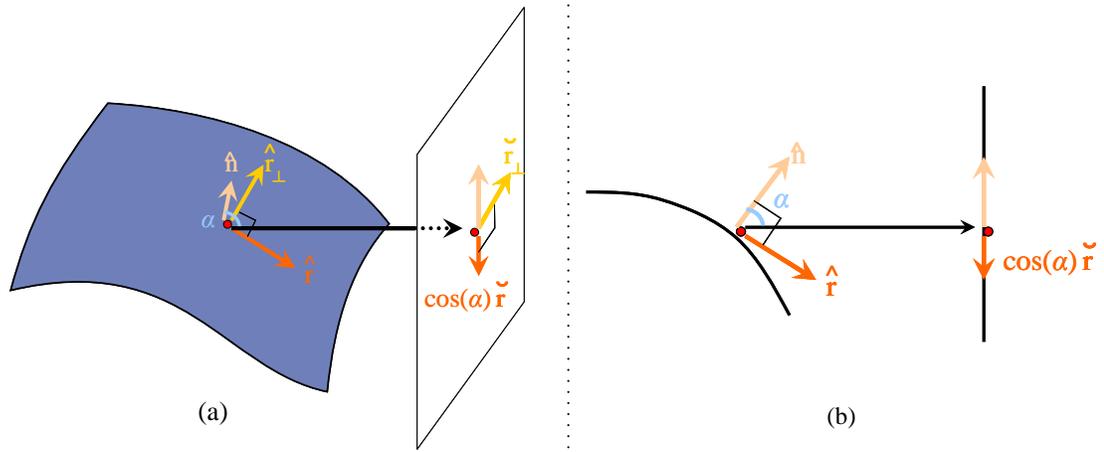


Figure 5.3: Orthographic projection onto the image screen. (a) shows a slightly tilted top-down view. (b) shows a slicing of the world with the plane that includes the normal and radial direction.

As seen in Figure 5.3, the two sets of radial bases (one normalized with the surface metric, one with the screen metric) are related as follows (when we think of screen-orthographic projection as a chart):

$$\hat{r}^a = \check{r}^a \cos \alpha, \quad (5.45)$$

$$\hat{r}_\perp^a = \check{r}_\perp^a. \quad (5.46)$$

With this information, we can define an orthographic projection to the screen (as opposed to the surface's tangent plane) as a mapping that sends a general vector in \mathbb{R}^3 to its image on the camera's screen:

$$A\hat{r}^a + B\hat{r}_\perp^a + C\hat{n}^a \mapsto (A \cos \alpha - C \sin \alpha)\check{r}^a + B\check{r}_\perp^a \quad (5.47)$$

This action can be described by a tensor Q_a^b that takes in a vector and spits out a new

vector:

$$Q_a{}^b := (\hat{r}_a - \tan \alpha \hat{n}_a) \hat{r}^b + \hat{r}_{\perp a} \hat{r}_{\perp}^b \quad (5.48)$$

Note that this projector is not symmetric, so the order of the indices matters. We also have the corresponding projector for covectors:

$$Q^a{}_b = (\hat{r}^a - \tan \alpha \hat{n}^a) \hat{r}_b + \hat{r}_{\perp}{}^a \hat{r}_{\perp b} \quad (5.49)$$

Since the projector is not symmetric, it is important that we use $Q^a{}_b$ to map covectors, not $Q_b{}^a$.

We can use the screen projection operators to define the screen's metric in terms of the surface's radial frame as well:

$$\check{g}_{ab} := Q_a{}^c Q_b{}^d ((\check{\vartheta}^1)_c (\check{\vartheta}^1)_d + \check{r}_{\perp c} \check{r}_{\perp d}) \quad (5.50)$$

$$= Q_a{}^c Q_b{}^d (\check{g}_{ce} \check{r}^e \check{g}_{df} \check{r}^f + \check{r}_{\perp c} \check{r}_{\perp d}) \quad (5.51)$$

$$= \cos^2 \alpha (\hat{r}_a - \tan \alpha \hat{n}_a) (\hat{r}_b - \tan \alpha \hat{n}_b) + \hat{r}_{\perp a} \hat{r}_{\perp b} \quad (5.52)$$

$$\check{g}^{ab} := \check{r}^a \check{r}^b + \check{r}_{\perp}{}^a \check{r}_{\perp}^b \quad (5.53)$$

$$= \frac{1}{\cos^2 \alpha} \hat{r}^a \hat{r}^b + \hat{r}_{\perp}{}^a \hat{r}_{\perp}^b \quad (5.54)$$

Note that the metric \check{g}_{ab} makes sense on general vectors in \mathbb{R}^3 , while \check{g}^{ab} only really makes sense for screen/surface vectors. Given Equations (5.45) and (5.46), it is easy to check that the pulling these tensors back to the screen's domain yields the previously defined screen tensors $\check{g}_{\bar{a}\bar{b}}$ and $\check{g}^{\bar{a}\bar{b}}$.

Because we have two metrics on the screen, care must be taken to not mix up notation. In our usage, if the same variable name is adorned with different patterns of lower and upper indices, we will maintain the following rule: lower indices will always be dual to upper indices via the surface metric, never the screen metric.

The full set of relations between the radial bases and their dual bases can now be stated:

$$\check{r}^a = \frac{1}{\cos \alpha} \hat{r}^a, \quad (5.55)$$

$$(\check{\vartheta}^1)_a := \check{g}_{ab} \check{r}^b \quad (5.56)$$

$$= \cos \alpha \hat{r}_a, \quad (5.57)$$

$$\check{r}_\perp^a = \hat{r}_\perp^a, \quad (5.58)$$

$$(\check{\vartheta}^2)_a := \check{g}_{ab} \check{r}_\perp^b \quad (5.59)$$

$$= \check{r}_{\perp a} \quad (5.60)$$

$$= \hat{r}_{\perp a}. \quad (5.61)$$

We have defined $(\check{\vartheta}^\mu)_a$ to be dual to the unit screen basis via the screen metric. This new notation is necessary in general since $(\check{\vartheta}^1)_a \neq \check{r}_a$. The former is dual to \check{r}^a with respect to the screen metric, and the latter is dual with respect to the surface metric. Since $(\check{\vartheta}^2)_a = \check{r}_{\perp a}$, we will sometimes use $\check{r}_{\perp a}$ since it is perhaps easier to parse.

The tensor Q_a^b , along with the covariant derivative for \mathbb{R}^3 , can be used to compute the screen-metric-compatible covariant derivative just as the tensor P_a^b (that projected vectors onto the tangent plane) can be used to compute the surface-metric-compatible co-

variant derivative. Just as the surface covariant derivative is known to be the \mathbb{R}^3 covariant derivative followed by projections onto the tangent plane, the screen metric-compatible covariant derivative is the 3D covariant derivative followed by screen-orthographic projections (applications of Q_a^b and Q^a_b). We can verify that this is the unique screen-metric-compatible covariant derivative by explicitly checking that such a construction obeys the metric compatibility requirements:

$$(\nabla_c \check{g}_{cd}) Q^c_a Q^d_b = 0_{cab} = \check{\nabla}_c \check{g}_{ab}, \quad (5.62)$$

$$(\nabla_c \check{g}^{cd}) Q_c^a Q_d^b = 0_c^{ab} = \check{\nabla}_c \check{g}^{ab}. \quad (5.63)$$

The left-hand equalities are checked via tedious calculations; the right-hand equalities follow from the definition of metric-compatible covariant derivative.

With the screen derivative now defined in terms of three-dimensional quantities, we can in theory compute any object defined on the screen (e.g., those directly detectable in the image) in terms of three-dimensional geometric information that is expressed with respect to the radial frame. If we furthermore choose a particular basis for the image screen, then tensor components can sometimes be computed even more easily. In a basis, components can often be computed as directional derivatives of some quantity; these computations can therefore leverage the fact that all derivatives essentially act the same on scalar functions. The main difficulty in this approach comes from defining the vector field directions (and their derivatives) along which one takes the derivatives of the scalar functions. We will pursue this line of analysis in the following paragraphs.

In an image of an illuminated surface, two “preferred directions” naturally arise

in the image. These are the gradient and isophote directions (on the screen). It is not surprising that many detectable image features such as Canny edges and Saint-Venant curves are defined using these quantities since they characterize the extremal changes in the image. The components of the illumination gradient in the screen's radial basis can be computed as follows:

$$\check{r}^a \nabla_a \cos \theta = \frac{1}{\cos \theta} \hat{r}^a (-\sin \theta (\kappa_r \hat{r}_a + \tau_r \hat{r}_{\perp a})) \quad (5.64)$$

$$= -\tan \theta \kappa_r, \quad (5.65)$$

$$\check{r}_{\perp}^a \nabla_a \cos \theta = \hat{r}_{\perp}^a (-\sin \theta (\kappa_r \hat{r}_a + \tau_r \hat{r}_{\perp a})) \quad (5.66)$$

$$= -\sin \theta \tau_r. \quad (5.67)$$

The illumination gradient (measured on screen as opposed to on surface) and the illumination isophote are then given by

$$\bar{\xi}^a = -\kappa_r \tan \theta \check{r}^a - \tau_r \sin \theta \check{r}_{\perp}^a \quad (5.68)$$

$$= -\frac{\kappa_r \sin \theta}{\cos^2 \theta} \hat{r}^a - \tau_r \sin \theta \hat{r}_{\perp}^a, \quad (5.69)$$

$$\bar{\xi}_{\perp}^a = \tau_r \sin \theta \check{r}^a - \kappa_r \tan \theta \check{r}_{\perp}^a \quad (5.70)$$

$$= \tau_r \tan \theta \hat{r}^a - \kappa_r \tan \theta \hat{r}_{\perp}^a. \quad (5.71)$$

Note that the screen illumination gradient direction is *not* the same as the surface illumination gradient direction, whereas the two isophote directions must of course match. We leave the bar on top of the vectors to remind us that these are gradients as computed on the screen domain using the screen metric. We denote the screen-normalized gradient and

isophote vector directions as $\check{\xi}^a$ and $\check{\xi}_\perp^a$:

$$\check{\xi}^a = -\frac{1}{\gamma} \left(\frac{\kappa_r}{\cos \theta} \hat{r}^a + \tau_r \cos \theta \hat{r}_\perp^a \right), \quad (5.72)$$

$$\check{\xi}_\perp^a = \frac{1}{\gamma} (\tau_r \hat{r}^a - \kappa_r \hat{r}_\perp^a), \quad (5.73)$$

where $\gamma := \sqrt{\kappa_r^2 + \tau_r^2 \cos^2 \theta}$.

We have expressed these vectors with respect to the screen's radial basis; however, such a basis is not directly detectable in an image, making it a somewhat arbitrary choice. Since the gradient and isophote directions are directly detectable in the image and are orthogonal on the screen, away from critical points of illumination, it is natural to express quantities in this "gradient basis." In particular, let us compute the components of the illumination hessian (as computed using the screen metric) in this basis.

In other words, at each point on screen, we would like to set up a set of Cartesian coordinates where our "x" direction aligns with the illumination gradient direction (on screen) and the "y" direction aligns with the illumination isophote direction. The second derivatives of the illumination hessian can then be computed by taking two successive derivatives in these Cartesian coordinate directions. The first derivative is easy to take since Equations (5.72) and (5.73) give us the gradient and isophote directions. However, taking the second derivatives requires more care. We cannot take the second derivative in the "x" direction simply by taking the directional derivative in the gradient direction twice. This is because the gradient vector field does not form a "straight line" vector field on the screen. The gradient direction rotates around on screen as one moves away from the image point of

interest. It therefore does not provide a nice Cartesian coordinate system in the image.³ We must therefore instead use a vector field “x” that evaluates to the gradient direction at the point of interest, but differs from the gradient vector field away from the point (differing in such a way as to not bend in the image). This implies that at each point on the screen, we must use a completely different “x” and “y” vector field. We will call these vector fields \check{x}^a and \check{y}^a (with the implicit understanding that at every point we have a completely different set of vector fields). Details for evaluating \check{x}^a and \check{y}^a and their derivatives are given in Appendix D.2.

The screen hessian (in the basis of screen unit gradient and isophote directions⁴) is given by:

$$\check{H}_{\check{\mu}}^{\check{\nu}} = \begin{bmatrix} \check{x}^{\check{a}} \check{x}^{\check{b}} \check{\nabla}_{\check{a}} \check{\nabla}_{\check{b}} \cos \theta & \check{y}^{\check{a}} \check{x}^{\check{b}} \check{\nabla}_{\check{a}} \check{\nabla}_{\check{b}} \cos \theta \\ \check{x}^{\check{a}} \check{y}^{\check{b}} \check{\nabla}_{\check{a}} \check{\nabla}_{\check{b}} \cos \theta & \check{y}^{\check{a}} \check{y}^{\check{b}} \check{\nabla}_{\check{a}} \check{\nabla}_{\check{b}} \cos \theta \end{bmatrix} \quad (5.74)$$

$$= \begin{bmatrix} \check{x}^{\check{a}} \nabla_a (\check{x}^{\check{b}} \nabla_b \cos \theta) & \check{y}^{\check{a}} \nabla_a (\check{x}^{\check{b}} \nabla_b \cos \theta) \\ \check{x}^{\check{a}} \nabla_a (\check{y}^{\check{b}} \nabla_b \cos \theta) & \check{y}^{\check{a}} \nabla_a (\check{y}^{\check{b}} \nabla_b \cos \theta) \end{bmatrix} \quad (5.75)$$

The first element on the diagonal is the Canny energy (expression that must go to zero for a point on a Canny edge). The off-diagonal terms are equal (since the illumination hessian is symmetric) and is equal to the Saint-Venant energy (the expression that must go to zero for a point on a Saint-Venant curve).⁵ This is confirmed by the fact that when a point is on a Saint-Venant curve, the gradient and isophote directions are eigenvectors, so the

³The vector field must not twist around on the screen for directional derivatives to match those computed using the usual partial derivatives in the image.

⁴This necessitates that we are not at a critical point of illumination.

⁵Technically, this is the Saint-Venant energy multiplied by a normalization scale factor $1/\gamma$.

off-diagonal terms must be zero in the gradient basis. In such a case, the diagonal terms give the eigenvalues corresponding to the gradient and isophote directions. In general, the components are given by the following:

$$\begin{aligned} \check{x}^a \nabla_a (\check{x}^b \nabla_b \cos \theta) &= -\frac{\tan \theta}{\gamma^2 \cos \theta} (\kappa_r \hat{r}^a + \tau_r \cos^2 \theta \hat{r}_\perp^a) (\kappa_r \nabla_a \kappa_r + \tau_r \cos^2 \theta \nabla_a \tau_r) \\ &\quad - \frac{1}{\cos^3 \theta} (\kappa_r^2 + \tau_r^2 \cos^4 \theta) \end{aligned} \quad (5.76)$$

$$\check{y}^a \nabla_a (\check{y}^b \nabla_b \cos \theta) = \frac{\sin \theta}{\gamma^2} (\tau_r \hat{r}^a - \kappa_r \hat{r}_\perp^a) (-\tau_r \nabla_a \kappa_r + \kappa_r \nabla_a \tau_r) - \frac{K}{\cos \theta} \quad (5.77)$$

$$\check{x}^a \nabla_a (\check{y}^b \nabla_b \cos \theta) = \check{y}^a \nabla_a (\check{x}^b \nabla_b \cos \theta) \quad (5.78)$$

$$= \check{x}^a \check{y}^b \nabla_a \nabla_b \cos \theta - \frac{K \kappa_r \tau_r \sin^2 \theta}{\gamma} \quad (5.79)$$

$$= \frac{\tan \theta}{\gamma^2} (\tau_r \hat{r}^a - \kappa_r \hat{r}_\perp^a) (\kappa_r \nabla_a \kappa_r + \tau_r \cos^2 \theta \nabla_a \tau_r). \quad (5.80)$$

5.3.2 Critical Points of Illumination

Here we state results that relate some of the curves at critical points of illumination. For these results, we make the following restrictions:

- (1) Directional lighting and orthographic camera: both the light position and view camera pinhole are located at points at infinity.
- (2) Miner's headlamp: The light position is co-located with the view camera's pinhole ($\theta = \alpha$ and $\phi = 0$).
- (3) Critical points not coming from normal incidence: It has been pointed out in [35] that—under directional lighting and Lambertian shading of generic surfaces—the generic

critical points of illumination either result from the light shining directly down the surface normal (resulting in isolated maxima) or cling to the surface's parabolic curves. We will only be interested in the latter case. This case is interesting because it ties the structural behavior of illumination to geometric properties of the surface.

- (4) Full curve zero set: Each of the curves we presented was defined as the zero set of some implicit function with culling of the solution space via some inequality relations. For simplicity, in this analysis, we consider each implicit function's entire zero set and ignore the inequality constraints.

Claim 1. *(Intersection) Under restrictions (1)-(4), parabolic lines, suggestive contours, and Saint-Venant curves all intersect at critical points of illumination.*

Claim 2. *(Tangency) Under restrictions (1)-(4), the suggestive contour tangent lines up with parabolic curve tangent at critical points of illumination.*

We now seek a condition under which Saint-Venant curves will also be tangent at the minimum of illumination.

Claim 3. *Under restrictions (1)-(4), the tangent direction of Saint-Venant curves correspond to eigenvector directions of the illumination hessian, $\check{H}_a^{\bar{b}}$, at critical points of illumination.*

Claim 4. *Under restrictions (1)-(4), the Saint-Venant curves are tangent to the suggestive and parabolic curves at critical points of illumination if and only if they all run perpendicular to the flat direction.*

The flat direction is simply the direction in which the normal curvature is zero. We know there is one such direction because the point lies on the parabolic curve.

As an aside, we can also relate suggestive contours to mean curvature curves:

Claim 5. *Under restrictions (1)-(4), a suggestive contour point falls on a minimal curve if and only if the radial direction is 45 degrees (on the surface) from a principal direction.*

5.3.3 Saint-Venant and Suggestive Energies

Away from critical points of illumination, it becomes more difficult to characterize the commonality in the behavior of the curves. One way to get some sense of the correspondence is to treat the implicit functions defining each curve as an “energy” that provides an ordinal measure of how far points are from the zero set. Since the suggestive contour energy is simply κ_r , we can easily compute the Saint-Venant energy on a suggestive contour (we make the same assumptions about light and camera configurations as in Section 5.3.2).

Away from critical points of illumination (i.e., away from points where both κ_r and τ_r are zero), we compute a variant of the Saint-Venant energy on a suggestive contour in which the isophote direction (the direction in which we look for gradient magnitude extrema) is

normalized to unit length on screen by plugging $\kappa_r = 0$ into Equation (5.80). This energy is given by

$$\hat{r}[\tau_r] \cdot \tan \theta . \quad (5.81)$$

This implies that if $\hat{r}[\tau_r]$ is very small, the Saint-Venant curves and suggestive contours nearly agree.

If the Saint-Venant energy is zero (so the two curves intersect), we can easily compute the eigenvalues of the illumination hessian as well. The eigenvalue associated with the isophote direction is given by

$$-\frac{\tan \theta}{\cos \theta} \hat{r}[\kappa_r] - \frac{K}{\cos \theta} . \quad (5.82)$$

Both terms are positive, with $\hat{r}[\kappa_r]$ strongly negative for a “strong” suggestive contour. The eigenvalue associated with the gradient direction is given by

$$-\hat{r}_\perp[\tau_r] \cdot \sin \theta + K \cos \theta . \quad (5.83)$$

The eigenvalue inequality condition for a Saint-Venant curve (which states that the isophote’s eigenvalue is larger in magnitude than the gradient’s eigenvalue) is then easily satisfied unless $\hat{r}_\perp[\tau_r]$ is very large in magnitude (on the order of $1/\cos^2 \theta$ times greater than the magnitude of $\hat{r}[\kappa_r]$).

This analysis suggests that the two types of curves closely match if the gradient of τ_r is small. (Recall that τ_r is a measure of how much the radial direction deviates from a

principal direction.)

5.3.4 Suggestive Contours and Shading

We now present some further connections between suggestive contours and shading in an image. It turns out that a lot of geometry can be deduced once one is told where in a shaded image the suggestive contours occur.

On suggestive contours, we have the identity $\kappa_r = 0$. Plugging this into the expression for the image's illumination gradient (given by Equation (5.68)), we find that the magnitude of the gradient on a suggestive contour is simply

$$\|\bar{\xi}\|_{\bar{g}} = \sqrt{-K} \sin \theta. \quad (5.84)$$

Since the image's shading gives us $\cos \theta$, this means that we can use the gradient magnitude to easily compute the Gaussian curvature of the surface. Furthermore, on a suggestive contour, the isophote direction is the same as the projected radial direction—which is also the projected normal direction. Given the projected normal direction and the angle between the normal and view direction (given by the shading $\cos \theta$), we can derive the three-dimensional unit normal to the surface. There is then enough information to integrate along the suggestive contours in the image and recover the *three-dimensional* suggestive contours!

From the diagonal entries of the illumination hessian in the gradient basis (Equations (5.76) and (5.77)), we can also compute third order derivative information about the

surface. Plugging in $\kappa_r = 0$, we can compute $\hat{r}_\perp[\tau_r]$ and $\hat{r}[\kappa_r]$. The quantity $\hat{r}[\kappa_r]$ is a measure of the strength of the suggestive contour. Recall that it is also the value that is used to cull the suggestive contours.

5.4 Discussion

We have presented a more analysis-friendly methodology for relating image information to surface information and applied these techniques to proving some results for suggestive contours and Saint-Venant curves. Analysis is made more intuitive because relational expressions are written in terms of curvatures and other important geometric quantities. Our results characterize curve behavior at critical points of illumination and point out other deducible facts for suggestive contours.

A more complete characterization of curves away from critical points of illumination is still lacking, and we hope the techniques presented here will be helpful in finishing such an analysis. Appendix D.4-D.6 discuss other relations that are also more generally applicable and may prove useful when pursuing such directions. These describe connections to principal directions, apparent principal directions, apparent curvatures, and general lighting conditions.

Chapter 6

Conclusions and Future Work

In this thesis, we described progress on a couple issues in perceptual image processing. We started with the problem of modeling color constancy and characterized the applicability of the generalized von Kries models in terms of rank constraints on the measured world. We then provided evidence that our world, as measured by some spectral databases, does indeed approximately meet these rank conditions and presented an algorithm for computing an optimal color basis for generalized von Kries modeling. These color constancy notions were then used to derive a new color space for illumination-invariant image processing. The derived color space also provides other useful perceptual features such as approximating perceptual distance with Euclidean distance and possessing an interpretation in terms of color opponent channels. Finally, we drew some connections between curves in an image and shape information.

These results also suggest a rich set of possible future directions to pursue. We

list some of these below according to their topic category. In choosing these directions, we do not limit ourselves to purely computer science questions.

Color constancy

It is known that humans do not possess perfect color constancy [18]. Likewise, our color space does not provide perfect color constancy (since the measurement tensor for the world is not exactly rank three). It would be interesting to see whether the deviations from perfect color constancy in humans can be predicted using our color space. In other words, do the deviations from perfect color constancy in both cases match?

Color space for illumination-invariant image processing

It would also be interesting to see whether the color space we propose for image processing can be used to predict other perceptual effects, either qualitatively or quantitatively. For instance, Appendix B.2 proves that our color space reproduces the well-known Weber's Law for brightness perception. This may explain the recent findings in [3] where human subjects determined that simple scalings of the linear color stimulus values give better results than more complicated tone-mapping operators when converting between high dynamic-range images and low dynamic-range images. Weber's Law states that—with regards to brightness—people are only sensitive to color ratios (in a linear color space) and so uniformly scaling each color component preserves the metric relations between colors. Our color space generalizes this notion to chromatic relighting as well.

Figure 3.1 also shows that the opponent colors in our color space are not exactly

the complementary colors. In XYZ space (which is a linear color space), the “chromaticity” geodesics curve slightly in the xy chromaticity plane. In fact, they curve in such a way that the opponent colors are biased toward the “unique hues” as opposed to strictly complementary colors.

Our color space can also be easily integrated into fibered space and measure theoretic approaches to modeling images. Such models are discussed in [4, 13, 27, 36]. Fibered space models nicely separate the spatial dimensions of the image (which are physical) from the color dimensions (which are, instead, perceptual). Our color space simply endows this topological structure with an appropriate metric. The measure theoretic models are useful for solving free-boundary problems, or ones in which the discontinuities in the image are to be solved for. Our color space, being a Euclidean space, allows all the analysis based on Lebesgue measure to carry over without change.

Shape from curves

We would like to also extend the techniques in Chapter 5 to give a more complete analysis of curves away from critical points of illumination. Such analyses will likely require some additional tools for characterizing global behaviors as well as local ones.

Bibliography

- [1] Janos Aczel, Jean-Claude Falmagne, and R. Duncan Luce. Functional equations in the behavioral sciences. *Math. Japonica*, 52(3):469–512, 2000.
- [2] Yair Adato, Yuriy Vasilyev, Ohad Ben-Shahar, and Todd Zickler. Toward a Theory of Shape from Specular Flow. *ICCV*, 2007.
- [3] Ahmet Oguz Akyuz, Erik Reinhard, Roland Fleming, Bernhard Riecke, and Heinrich Buelthoff. Do hdr displays support ldr content? a psychophysical evaluation. *SIGGRAPH*, 26(3), 2007.
- [4] Gilles Aubert and Pierre Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer, 2nd edition, 2006.
- [5] R.K. Bajcsy and L.I. Lieberman. Texture gradient as a depth cue. *CGIP*, 5:52–67, 1976.
- [6] K. Barnard, F. Ciurea, and B. Funt. Sensor sharpening for computational color constancy. *JOSA A*, 18(9):2728–2743, 2001.
- [7] Kobus Barnard, Lindsay Martin, Brian Funt, and Adam Coath. A data set for colour research. *Color Research and Application*, 27(3):147–151, 2002.
- [8] Roy S. Berns, David H. Alman, Lisa Reniff, Gregory D. Snyder, and Mitchell R. Balonon-Rosen. Visual determination of suprathreshold color-difference tolerances using probit analysis. *Color Research and Applications*, 16(5):297–316, 1991.
- [9] Manfredo P. Do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.
- [10] Sean M. Carroll. *An Introduction to General Relativity: Spacetime and Geometry*. Addison-Wesley, 2004.

- [11] Enrique Castillo and Maria Reyes Ruiz-Cobo. *Functional Equations and Modelling in Science and Engineering*. CRC, 1992.
- [12] Andrew N. Chalmers, Snjezana Soltic, and Radhika Jammalamadaka. A case for a ciecam02 colour appearance space. *Manukau Dept. ECE Conf. Papers*, 2007.
- [13] Tony F. Chan and Jianhong Shen. *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*. SIAM, 2005.
- [14] Hamilton Y. Chong, Steven J. Gortler, and Todd Zickler. The von kries hypothesis and a basis for color constancy. *ICCV*, 2007.
- [15] Doug DeCarlo, Adam Finkelstein, Szymon Rusinkiewicz, and Anthony Santella. Suggestive contours for conveying shape. *SIGGRAPH*, 22(3):848–855, July 2003.
- [16] M. S. Drew and G. D. Finlayson. Spectral sharpening with positivity. *JOSA A*, 17(8):1361–1370, 2000.
- [17] Fritz Ebner and Mark D. Fairchild. Development and testing of a color space (ipt) with improved hue uniformity. *Proc. 6th Color Imaging Conf.*, pages 8–13, 1998.
- [18] Mark D. Fairchild. *Color Appearance Models*. Wiley-IS&T, 2nd edition, 2005.
- [19] Gustav Theodor Fechner. *In Sachen der Psychophysik*. Leipzig, 1877.
- [20] G. D. Finlayson, S. D. Hordley, and A. Alsam. Investigating von kries-like adaptation using local linear models. *Color Research and Application*, 31(2):90–101, 2006.
- [21] G.D. Finlayson, M.S. Drew, and B.V. Funt. Diagonal transforms suffice for color constancy. In *ICCV93*, pages 164–171, 1993.
- [22] G.D. Finlayson, M.S. Drew, and B.V. Funt. Enhancing von kries adaptation via sensor transformations. *SPIE*, 1993.
- [23] G.D. Finlayson, M.S. Drew, and B.V. Funt. Spectral sharpening: sensor transformations for improved color constancy. *JOSA A*, 11(5):1553–1563, 1994.
- [24] D. A. Forsyth. A novel algorithm for color constancy. *IJCV*, 5(1):5–35, 1990.
- [25] D. H. Foster, S. M. C. Nascimento, and K. Amano. Information limits on neural identification of coloured surfaces in natural scenes. *Visual Neuroscience*, 21:331–336, 2004.
- [26] Jean Gallier. *Geometric Methods and Applications: For Computer Science and Engineering*. Springer, 2001.

- [27] T. Georgiev. Covariant derivatives and vision. In *ECCV '06*, pages Part IV, 56–69. Springer-Verlag, 2006.
- [28] R. A. Harshman. Foundations of the parafac procedure: Models and conditions for an “explanatory” multimodal factor analysis. *UCLA Working Papers in Phonetics*, 16:1–84, 1970.
- [29] B.K.P. Horn. Shape from shading: A method for obtaining the shape of a smooth opaque object from one view. *MIT PhD Thesis*, 1970.
- [30] Jürgen Jost. *Riemannian Geometry and Geometric Analysis*. Springer, 2000.
- [31] Deane B. Judd and Gunter Wyszecki. *Color In Business, Science, and Industry*. John Wiley and Sons, 1975.
- [32] Tilke Judd, Frédo Durand, and Edward H. Adelson. Apparent ridges for line drawing. *SIGGRAPH*, 26(3), 2007.
- [33] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and K. Sobel. A variational framework for retinex. *IJCV*, 2002.
- [34] Jan J. Koenderink. What does the occluding contour tell us about solid shape? *Perception*, 13(3):321–330, 1984.
- [35] Jan J. Koenderink and Andrea J. van Doorn. Illuminance critical points on generic smooth surfaces. *J. Opt. Soc. Am. A*, 10(5), 1993.
- [36] Jan J. Koenderink and Andrea J. van Doorn. Image processing done right. In *ECCV '02*, pages 158–172. Springer-Verlag, 2002.
- [37] Lieven De Lathauwer, Bard De Moor, and Joos Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix Analysis*, 21(4):1253–1278, 2000.
- [38] M.I.A. Lourakis. levmar: Levenberg-marquardt nonlinear least squares algorithms in C/C++. <http://www.ics.forth.gr/~lourakis/levmar/>. [Accessed Dec. 2007.].
- [39] R. Duncan Luce. Let’s not promulgate either fechner’s erroneous algorithm or his unidimensional approach. *Behavior and Brain Sciences*, 16:155–156, 1993.
- [40] M. Rinnier Luo. Colour difference formulae: Past, present and future. *CIE Expert Symposium*, 2006.
- [41] L.T. Maloney. *Color Vision: From Genes to Perception*, chapter Physics-based approaches to modeling surface color perception, pages 387–422. Cambridge University Press, 1999.

- [42] D. Marimont and B.A. Wandell. Linear models of surface and illuminant spectra. *JOSA A*, 9:1905–1913, 1992.
- [43] J. J. McCann. Do humans discount the illuminant? *Proc. SPIE*, 5666:9–16, 2005.
- [44] Nathan Moroney, Mark D. Fairchild, Robert W. G. Hunt, Changjun Li, M. Ronnier Luo, and Todd Newman. The ciecam02 color appearance model. In *Color Imaging Conference*, pages 23–27, 2002.
- [45] U. of Joensuu Color Research Spectral Database. <http://spectral.joensuu.fi/>.
- [46] Stephen E. Palmer. *Vision Science: Photons to Phenomenology*. MIT Press, 1999.
- [47] G. Papandreou and P. Maragos. Multigrid geometric active contour models. *IEEE Trans. on Image Processing*, 16(1):229–240, January 2007.
- [48] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. *SIG-GRAPH*, 22(3):313–318, 2003.
- [49] H. L. Resnikoff. Differential geometry and color perception. *Journal of Mathematical Biology*, 2(4):97–131, 1975.
- [50] H. L. Resnikoff. *The Illusion of Reality*. Springer-Verlag, 1989.
- [51] Gaurav Sharma, Wencheng Wu, and Edul N. Dalal. The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research and Application*, 30(1):21–30, 2005.
- [52] N.D. Sidiropoulos. Low-rank decomposition of multi-way arrays: A signal processing perspective. In *IEEE SAM*, July 2004.
- [53] Michael Spivak. *A Comprehensive Introduction to Differential Geometry*. Publish or Perish, 3rd edition, 1999.
- [54] J. van de Weijer and T. Gevers. Color constancy based on the grey-edge hypothesis. *ICIP*, 2, 2005.
- [55] JJ Vos. From lower to higher colour metrics: a historical account. *Clinical & Experimental Optometry*, 86(6):348–360, 2006.
- [56] Robert M. Wald. *General Relativity*. University of Chicago Press, 1984.
- [57] Michael A. Webster. Human colour perception and its adaptation. *Network: Computation in Neural Systems*, 7, 1996.

-
- [58] Gerhard West and Michael H. Brill. Necessary and sufficient conditions for von kries chromatic adaptation to give color constancy. *Journal of Mathematical Biology*, 15(2):249–258, 1982.
- [59] Klaus Witt. Geometric relations between scales of small colour differences. *Color Research and Applications*, 24(2):78–92, 1999.
- [60] G. Wyszecki and W.S. Stiles. *Color Science: concepts and methods, quantitative data and formulae*. Wiley, 1982.

Appendix A

Color Constancy Proofs

A.1 Conditions for Success

For simplicity, we only consider generic conditions.

A.1.1 Proof of Proposition 1

We first show that if a measurement tensor supports doubly linear color constancy, then it can be transformed into the core tensor form of Figure 2.2 via changes in illuminant and material bases.

Lemma 2. *If $\forall j, j', \exists \bar{D}(R_j, R_{j'})$ satisfying the definition of linear relational color constancy, then \exists a single change of basis for the space of illuminants, \mathcal{E} , such that $\forall j \in \{1, \dots, J\}$ the $3 \times I$ matrix $\Omega_{(j)}$ is zero past the third column.*

Proof. (Lemma 2) By definition of linear relational color constancy, we have $\bar{D}(R_j, R_{j'})$ is linear, and $\forall E \in \mathcal{E}$, its corresponding coordinate vector (in the basis of illuminant spectra used to create the measurement tensor) \vec{E} satisfies $\Omega_{(j')} \vec{E} = \bar{D}(R_j, R_{j'}) \Omega_{(j)} \vec{E}$. Let $N_j \subseteq \mathcal{E}$ be the subset of \mathcal{E} whose coordinate vectors (also in the basis of illuminant spectra used to create the measurement tensor), $\{\vec{n}\}$, constitute the null space of $\Omega_{(j)}$. Since $\Omega_{(j)}$ is rank-3, N_j^\perp is 3 dimensional. Given the above, we have that $\forall \vec{n}$

$$\Omega_{(j')}(\vec{E} + \vec{n}) = \bar{D}(R_j, R_{j'}) \Omega_{(j)}(\vec{E} + \vec{n}) \quad (\text{A.1})$$

$$\Omega_{(j')} \vec{E} + \Omega_{(j')} \vec{n} = \bar{D}(R_j, R_{j'}) \Omega_{(j)} \vec{E} + 0 \quad (\text{A.2})$$

$$\Omega_{(j')} \vec{n} = 0 \quad (\text{A.3})$$

So $N_j \subseteq N_{j'}$. By symmetry under interchange of j and j' , $N_{j'} \subseteq N_j$, so $N_j = N_{j'} \forall j, j'$. Fix a j , and choose a basis with the first three vectors spanning N_j^\perp and orthogonal to N_j , and the rest of the basis vectors spanning N_j and orthogonal to N_j^\perp . Under this choice, every $\Omega_{(j)}$ will have zeros past the third column since the decompositions $\text{Ker}(\Omega_{(j)})^\perp \oplus \text{Ker}(\Omega_{(j)})$ are the same for all j . ■

By symmetry of Lemma 2, we have a similar statement for $D(E_i, E_{i'})$ and a change of basis for \mathcal{R} . Hence we can get the core tensor form via change of illuminant and material bases.

We now prove the converse direction. The ability to achieve the core tensor form of Figure 2.2 implies there is a single change of illuminant basis transforming the measurement tensor's $\Omega_{(j)}$ slices into the block form in Figure A.1. Denote the nonzero 3x3 sub-block of $\Omega_{(j)}$ as $\tilde{\Omega}_{(j)}$. Now define for every pair j, j' the linear operator $\tilde{\Omega}_{(j')} \tilde{\Omega}_{(j)}^{-1}$.

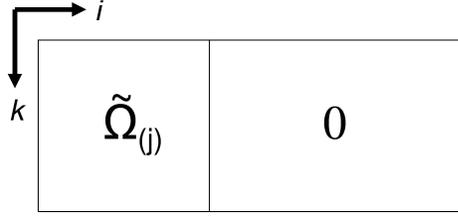


Figure A.1: Following a change of illuminant basis, the $3 \times I$ measurement tensor slice $\Omega_{(j)}$ is partitioned into a nonzero 3×3 sub-block $\tilde{\Omega}_{(j)}$ and a zero $3 \times (I - 3)$ sub-block as shown.

Observe that this operator maps the columns of matrix $\Omega_{(j)}$ (color vectors of material R_j) to the corresponding columns of $\Omega_{(j')}$ (color vectors of material $R_{j'}$). The nonzero columns are mapped appropriately by construction, and the zero columns are mapped to zero by linearity. This means that for all basis illuminant spectra of \mathcal{E} , the operator $\tilde{\Omega}_{(j')} \tilde{\Omega}_{(j)}^{-1}$ correctly maps the associated material R_j colors to the appropriate material $R_{j'}$ colors. Since color measurements are linear in the illuminant, and the colors associated with each basis illuminant spectrum are mapped correctly, we can conclude that $\tilde{\Omega}_{(j')} \tilde{\Omega}_{(j)}^{-1}$ acts as $\bar{D}(R_j, R_{j'})$ does on *all* color vectors. This implies:

$$\bar{D}(R_j, R_{j'}) = \tilde{\Omega}_{(j')} \tilde{\Omega}_{(j)}^{-1}. \quad (\text{A.4})$$

Hence, \bar{D} is linear. Since \bar{D} does not depend on illuminants, it is invariant to any illuminant basis change and so the initial change of illuminant basis did not alter whether the measurement tensor supports linear relational color constancy. Likewise, a change of material basis making all the $\Lambda_{(i)}$ slices null past the third column implies D is linear, and the tensor supports doubly linear color constancy.

A.1.2 Proof of Lemma 1

Our argument will proceed in two steps. We first prove a version of the lemma (referred to as Lemma 3) in which “measurement tensor” is replaced by “core tensor of the form in Figure 2.2”. We then prove Lemma 1 using this core tensor version (Lemma 3).

Lemma 3. *A core tensor of the form in Figure 2.2 supports generalized diagonal color constancy iff there exists a change of color basis such that core tensor slice $\Gamma^{(k)}$ is a rank-1 matrix $\forall k$.*

Proof. (Lemma 3) We first prove that if the core tensor has rank-1 $\Gamma^{(k)}$ slices under some change of color basis, then $\bar{D}(R_j, R_{j'})$ is diagonal $\forall j, j'$ (under that same color basis). This implies the core tensor supports generalized diagonal color constancy. By equation (A.4), $\bar{D}(R_j, R_{j'}) = \tilde{\Omega}_{(j')} \tilde{\Omega}_{(j)}^{-1}$. Denote the rows of slice $\tilde{\Omega}_{(j)}$ as $\vec{u}_1^T, \vec{u}_2^T, \vec{u}_3^T$. Since the $\Gamma^{(k)}$ slices are rank-1 matrices, the rows of another slice $\tilde{\Omega}_{(j')}$ are $\gamma_1 \vec{u}_1^T, \gamma_2 \vec{u}_2^T, \gamma_3 \vec{u}_3^T$ for some scales $\gamma_1, \gamma_2, \gamma_3$. Let \vec{u}_k^* denote the column vector that is dual to the row \vec{u}_k^T (i.e., $\vec{u}_k^T \vec{u}_{k'}^* = \delta_{kk'}$). We then have:

$$\bar{D}(R_j, R_{j'}) = \tilde{\Omega}_{(j')} \tilde{\Omega}_{(j)}^{-1} \quad (\text{A.5})$$

$$= \begin{bmatrix} \gamma_1 \vec{u}_1^T \\ \gamma_2 \vec{u}_2^T \\ \gamma_3 \vec{u}_3^T \end{bmatrix} \begin{bmatrix} \vec{u}_1^* & \vec{u}_2^* & \vec{u}_3^* \end{bmatrix} \quad (\text{A.6})$$

$$= \begin{bmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma_3 \end{bmatrix} \quad (\text{A.7})$$

Thus $\bar{D}(R_j, R_{j'})$ is diagonal.

We now prove that support for generalized diagonal color constancy implies the core tensor has rank-1 $\Gamma^{(k)}$ slices under an appropriate change of color basis. If the core tensor supports generalized diagonal color constancy, then there exists a change of color basis such that the linear map $\bar{D}(R_j, R_{j'}) := \tilde{\Omega}_{(j')} \tilde{\Omega}_{(j)}^{-1}$ is in fact diagonal $\forall j, j'$. It then suffices to show that $\bar{D}(R_j, R_{j'})$ being diagonal for all j, j' implies $\Gamma^{(k)}$ is a rank-1 matrix $\forall k$. Denote the rows of slice $\tilde{\Omega}_{(j)}$ as $\vec{u}_1^T, \vec{u}_2^T, \vec{u}_3^T$. As before, let $\vec{u}_1^*, \vec{u}_2^*, \vec{u}_3^*$ be their duals. Denote the rows of slice $\tilde{\Omega}_{(j')}$ as $\vec{v}_1^T, \vec{v}_2^T, \vec{v}_3^T$. Let $\gamma_1, \gamma_2, \gamma_3$ be the diagonal elements of $\bar{D}(R_j, R_{j'})$. From equation (A.4), $\bar{D}(R_j, R_{j'})$ being diagonal implies:

$$\vec{v}_k^T \vec{u}_{k'}^* = \gamma_k \delta_{kk'} \quad (\text{A.8})$$

The uniqueness of the dual relationship implies the row vectors \vec{u}_k^T and \vec{v}_k^T must be the same up to scale. So the $\Gamma^{(k)}$ slices are rank-1 matrices. ■

Proof. (Lemma 1) We will make use of the following 3 observations which are not hard to show: (1) changing material and illuminant bases does not affect whether the $\Gamma^{(k)}$ slices are rank-1 matrices; (2) changing illuminant, material, and color bases does not affect whether a tensor supports generalized diagonal color constancy; (3) the order in which one changes illuminant, material, and color bases does not matter – the resulting tensor is the same.

Forward direction: Support for generalized diagonal color constancy implies rank-1 $\Gamma^{(k)}$ in some color basis. A measurement tensor supporting generalized diagonal color constancy must also support doubly linear color constancy (which is a looser restriction). The tensor can then be reduced to the core tensor form of Figure 2.2 via changes in illuminant and material bases (and the core tensor must also support generalized diagonal color constancy). By the forward direction of Lemma 3, the core tensor has rank-1 $\Gamma^{(k)}$ slices under some color transform. This implies the original measurement tensor has rank-1 $\Gamma^{(k)}$ slices under the same color transform.

Backward direction: Rank-1 $\Gamma^{(k)}$ in some color basis implies support for generalized diagonal color constancy. First transform to the color basis in which the Γ^k slices are rank-1 matrices. We now show that a measurement tensor with rank-1 $\Gamma^{(k)} \forall k$ can be transformed into the core tensor form of Figure 2.2 via changes in illuminant and material bases; the resulting core tensor also has rank-1 $\Gamma^{(k)}$ slices by observation 1.

Without loss of generality we consider the $\Omega_{(j)}$ slices. The $\Lambda_{(i)}$ slices proof is symmetric. The following computation shows that any column i of the $3 \times I$ matrix $\Omega_{(j)}$ can be (in the generic case) written as a linear combination of the first 3 columns. Let \vec{u}_k^T be a row vector consisting of the first 3 elements of the k -th row of $\Omega_{(j)}$. Let v_k be the k -th

element of the i -th column of $\Omega_{(j)}$. We seek weights w_1, w_2, w_3 such that:

$$\begin{bmatrix} \vec{u}_1^T \\ \vec{u}_2^T \\ \vec{u}_3^T \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad (\text{A.9})$$

Since the leftmost matrix is a 3×3 invertible matrix, there exists a unique solution. Suppose we chose a different slice $\Omega_{(j')}$. The rows of this matrix are scales of the original $\Omega_{(j)}$ rows because $\Gamma^{(k)}$ is a rank-1 matrix $\forall k$. Let the scales be $\gamma_1, \gamma_2, \gamma_3$. The new system of equations to solve (with unknown weights w'_1, w'_2, w'_3) is:

$$\begin{bmatrix} \gamma_1 \vec{u}_1^T \\ \gamma_2 \vec{u}_2^T \\ \gamma_3 \vec{u}_3^T \end{bmatrix} \begin{bmatrix} w'_1 \\ w'_2 \\ w'_3 \end{bmatrix} = \begin{bmatrix} \gamma_1 v_1 \\ \gamma_2 v_2 \\ \gamma_3 v_3 \end{bmatrix} \quad (\text{A.10})$$

This is the same linear system as (A.9) and therefore possesses the same unique solution. We get that $w'_k = w_k$. If we replace the i -th illuminant spectrum $E_i(\lambda)$ with the spectrum $E_i(\lambda) - w_1 E_1(\lambda) - w_2 E_2(\lambda) - w_3 E_3(\lambda)$, then the transformation is an invertible change of basis (since the initial illuminant samples were linearly independent), the resulting set of illuminant spectra spans the same linear space, and the i -th column of $\Omega_{(j)}$ is now zero for all j . Repeating this for every $i > 3$ makes $\Omega_{(j)}$ zero past the third column for all j .

By symmetry of the argument for the $\Lambda_{(i)}$ slices, we reduce the measurement tensor to the core tensor form of Figure 2.2 via changes in illuminant and material bases. Note furthermore that the core tensor's $\Gamma^{(k)}$ slices must also be rank-1 matrices since the $\Gamma^{(k)}$ slices before changes in illuminant and material bases were. By the backward direction of Lemma 3, the core tensor supports generalized diagonal color constancy. This implies that the original measurement tensor also supports generalized diagonal color constancy. ■

A.2 The Space of Solutions

Here we further analyze the space of rank 3 tensors. Without loss of generality, we will concern ourselves with $3 \times 3 \times 3$ core tensors for simplicity. Let C be the color basis in which color constancy mappings are diagonal. Questions we would like to answer include the following: (1) What are the sets of possible slices $\tilde{\Omega}_{(j)}$ and $\tilde{\Lambda}_{(i)}$ that we can observe? (2) If some slices of the tensor are fixed (perhaps by mandating the inclusion of certain materials and lights in our world), what are the possible ways in which to fill in the rest of the measurement tensor?

To formalize these questions, we define the space of matrices that map colors to colors and are diagonalized in the color basis C : Let $\mathcal{M} := \{M \in \text{GL}_3 \mid C^{-1}MC \text{ is diagonal}\}$. We then want to determine the maximal sets $\Omega \subseteq \text{GL}_3$, such that for all $\tilde{\Omega}_{(j)}, \tilde{\Omega}_{(j')} \in \Omega$, the matrix $\tilde{\Omega}_{(j')} \tilde{\Omega}_{(j)}^{-1}$ is in \mathcal{M} . Likewise we want to determine the maximal sets $\Lambda \subseteq \text{GL}_3$, such that for all $\tilde{\Lambda}_{(i)}, \tilde{\Lambda}_{(i')} \in \Lambda$, the matrix $\tilde{\Lambda}_{(i')} \tilde{\Lambda}_{(i)}^{-1}$ is in \mathcal{M} .

Proposition 2. *The possible Ω and Λ are precisely the right cosets $\mathcal{M}A$, where $A \in \text{GL}_3$.*

Proof. (Proposition 2) We will prove the statement for Ω . The Λ case is completely analogous. We first show that every set Ω is contained in some right coset $\mathcal{M}A$, for some $A \in \text{GL}_3$. By definition of Ω , for every pair of slices $\tilde{\Omega}_{(j)}, \tilde{\Omega}_{(j')} \in \Omega$, we have that

$$\tilde{\Omega}_{(j')} \tilde{\Omega}_{(j)}^{-1} = M_{(j'j)} \text{ for some } M_{(j'j)} \in \mathcal{M}. \quad (\text{A.11})$$

This implies that every pair of slices $\tilde{\Omega}_{(j)}$ and $\tilde{\Omega}_{(j')}$ then differ by some $M_{(j'j)}$ on the left:

$$\tilde{\Omega}_{(j')} = M_{(j'j)} \tilde{\Omega}_{(j)}. \quad (\text{A.12})$$

Choose $A := \tilde{\Omega}_{(j)}$. Then every slice $\tilde{\Omega}_{(j')}$ is in $\mathcal{M}A$.

We now prove that every right coset $\mathcal{M}A$ is contained in some Ω (i.e., $\mathcal{M}A \subseteq \Omega$).

Let $M_{(1)}A$ and $M_{(2)}A$ be two elements in $\mathcal{M}A$. Then we have the following:

$$M_{(1)}A(M_{(2)}A)^{-1} = M_{(1)}AA^{-1}M_{(2)}^{-1} \quad (\text{A.13})$$

$$= M_{(1)}M_{(2)}^{-1} \in \mathcal{M}. \quad (\text{A.14})$$

This is true for every $M_{(1)}A$ and $M_{(2)}A$ in $\mathcal{M}A$, so these must be elements in some Ω since Ω is by definition a maximal set matrices obeying this property.

Finally, we know that two cosets are either equal or completely disjoint. Two cosets cannot have a proper and nontrivial intersection. (We can also use the fact that Ω is a maximal set, so it cannot properly contain a different Ω' .) Therefore, if Ω and $\mathcal{M}A$ have any intersection, $\Omega \subseteq \mathcal{M}A$ and $\mathcal{M}A \subseteq \Omega$. So they are exactly equal. ■

Appendix B

Color Space Proofs

B.1 Deriving the Functional Form

In the notation of Equation (3.3), define the displacement vector $V(\vec{x}, \vec{x}') := F(\vec{x}) - F(\vec{x}')$.

V satisfies the following property (reflecting the Euclidean nature of our space): $\forall \vec{x}, \vec{x}', \vec{x}''$,

$$V(\vec{x}'', \vec{x}') + V(\vec{x}', \vec{x}) = F(\vec{x}'') - F(\vec{x}') + F(\vec{x}') - F(\vec{x}) \quad (\text{B.1})$$

$$= V(\vec{x}'', \vec{x}). \quad (\text{B.2})$$

Meanwhile, Equation (3.3) states that $V(\vec{x}, \vec{x}')$ is unchanged if both arguments are transformed by CDC^{-1} . Define $\vec{y} := C^{-1}\vec{x}$ and $\vec{y}' := C^{-1}\vec{x}'$. Then since equation (3.3) holds for all diagonal matrices D , let us choose the diagonal matrix whose diagonal elements are the

reciprocal of the components of \vec{y}' . This gives us: $\forall \vec{x}, \vec{x}'$,

$$V(\vec{x}, \vec{x}') = V(CDC^{-1}[x_1, x_2, x_3]^T, CDC^{-1}[x'_1, x'_2, x'_3]^T) \quad (\text{B.3})$$

$$= V\left(C \begin{bmatrix} y_1 & y_2 & y_3 \\ y'_1 & y'_2 & y'_3 \end{bmatrix}^T, C[1, 1, 1]^T\right). \quad (\text{B.4})$$

Since C and $C[1, 1, 1]^T$ are just fixed constants, define:

$$W\left(\frac{y_1}{y'_1}, \frac{y_2}{y'_2}, \frac{y_3}{y'_3}\right) := V\left(C \begin{bmatrix} y_1 & y_2 & y_3 \\ y'_1 & y'_2 & y'_3 \end{bmatrix}^T, C[1, 1, 1]^T\right) \quad (\text{B.5})$$

The function $W(y_1/y'_1, y_2/y'_2, y_3/y'_3)$ inherits the additive property of Equations (B.1)-(B.2).

In other words:

$$W\left(\frac{y''_1}{y'_1}, \frac{y''_2}{y'_2}, \frac{y''_3}{y'_3}\right) + W\left(\frac{y'_1}{y_1}, \frac{y'_2}{y_2}, \frac{y'_3}{y_3}\right) = W\left(\frac{y''_1}{y_1}, \frac{y''_2}{y_2}, \frac{y''_3}{y_3}\right). \quad (\text{B.6})$$

If we make the variable substitutions $u_i = y''_i/y'_i$, $v_i = y'_i/y_i$, then the additive property becomes the set of functional equations:

$$W_i(u_1 v_1, u_2 v_2, u_3 v_3) = W_i(u_1, u_2, u_3) + W_i(v_1, v_2, v_3), \quad i = 1, 2, 3. \quad (\text{B.7})$$

The only non-constant and continuous solutions (defined for positive u_i and v_i) are (see [11]):

$$W_i(u_1, u_2, u_3) = \sum_{j=1}^3 a_{ij} \ln u_j, \quad a_{ij} \in \mathbb{R}. \quad (\text{B.8})$$

Substituting back into Equation (B.5) yields the parameterization F :

$$F(\vec{x}) = A(\hat{\ln}(C^{-1} \vec{x})) + \vec{k}, \quad (\text{B.9})$$

where A and C^{-1} are invertible 3×3 matrices and $\hat{\ln}$ denotes component-wise natural logarithm. The vector \vec{k} might as well be taken to be zero since the range of F is an affine space (only displacements, their magnitudes, and higher order descriptors have illumination-invariant meaning).

B.2 Recovering Weber's Law

Our derived metric recovers Weber's Law for brightness perception, which states that perceived brightness grows logarithmically with stimulus intensity.

For $\lambda > 0$:

$$d(\lambda \vec{x}, \vec{x}) = \|F(\lambda \vec{x}) - F(\vec{x})\| \quad (\text{B.10})$$

$$= \|A(\hat{\ln} C^{-1} \vec{x} + \hat{\ln}(\lambda \cdot \vec{1}) - \hat{\ln} C^{-1} \vec{x})\| \quad (\text{B.11})$$

$$= |\ln(\lambda)| \cdot \|A \cdot \vec{1}\| \quad (\text{B.12})$$

where $\vec{1}$ is a 3-vector of all 1's and $\|A \cdot \vec{1}\|$ is a positive constant.

Appendix C

More Differential Geometry

C.1 Non-coordinate Bases

As mentioned in Chapter 4, not all possible bases for a manifold arise as coordinate bases. In particular, a set of orthonormal basis vector fields (frame fields) defined over an open neighborhood of the surface cannot correspond to coordinate basis vector fields unless the manifold has no Gaussian curvature. This can be seen by noting the following: in a coordinate basis, Gaussian curvature can be computed by some function of the metric's components (and its derivatives) in that basis; for an orthonormal frame field, the metric components are constant and the derivatives are identically zero; therefore, in general, a manifold with nontrivial Gaussian curvature does not possess coordinate bases that form an orthonormal frame field.

Coordinate bases are natural in the sense that manifolds are defined using coordinate charts. However, there is no reason computations must be carried out in a coordinate basis. As intuition would suggest, geometric quantities are invariant more generally to *all* choices of basis fields (coordinate or noncoordinate). In fact, chosen bases do not have to be related to geometry at all. For our purposes, however, we will choose orthonormal bases (and when given hypersurfaces, they will be adapted to the hypersurface). These bases, although noncoordinate, are also limited in some sense by the intrinsic geometry of the manifold. As discussed in the other sections of this Appendix, the “turning” of these frames is related to the manifold’s Gaussian curvature.

For this section, we will adopt the convention of using the early Latin letters (e.g., a, b, c, \dots) for tensor placeholders, and both Greek letters ($\mu, \lambda, \sigma, \dots$) and later Latin letters starting from “i” onward (e.g., i, j, k, \dots) for quantities concerning particular choices of bases.

Let $\left\{ \left(\frac{\partial}{\partial x^\mu} \right)^a \right\}$ be a coordinate basis and $\{(e_i)^a\}$ be a noncoordinate basis. Their corresponding dual bases are $\{(dx^\mu)_a\}$ and $\{(\vartheta^i)_a\}$ respectively. At each point on the manifold, the bases are related by a matrix with components $e^j{}_\lambda$ (the name “ e ” is a standard choice):

$$\left(\frac{\partial}{\partial x^\lambda} \right)^a = e^j{}_\lambda (e_j)^a. \quad (\text{C.1})$$

The dual bases are then related by the inverse matrix, which we denote as $e_k{}^\mu$ (although unnecessarily confusing, this repetition of variable names for both the matrix and its inverse

is standard—the two matrices are distinguished only by the placement of their indices):

$$(dx^\mu)_a = e_k^\mu (\vartheta^k)_a. \quad (\text{C.2})$$

These same matrices can be used to transform vector components as well. For instance, a vector v^a with coordinates v^i in the $\{(e_i)^a\}$ basis can be reassigned components v^λ in the $\left\{\left(\frac{\partial}{\partial x^\lambda}\right)^a\right\}$ basis as follows:

$$v^\lambda = e_i^\lambda v^i. \quad (\text{C.3})$$

Given the rule for computing covariant derivatives in a coordinate basis, we can rewrite the expression in terms of a noncoordinate basis. By doing so, we see that computing the covariant derivative with respect to a noncoordinate basis is symbolically identical.

$$\nabla_a v^b = (\partial_\mu v^\lambda + \Gamma_\mu^\lambda{}^\sigma v^\sigma)(dx^\mu)_a \left(\frac{\partial}{\partial x^\lambda}\right)^b \quad (\text{C.4})$$

$$= (e_i^\lambda \partial_\mu v^i + v^i \partial_\mu e_i^\lambda + \Gamma_\mu^\lambda{}^\sigma e_i^\sigma v^i) e_k^\mu (\vartheta^k)_a e^j{}_\lambda (e_j)^b \quad (\text{C.5})$$

$$= \left(e_k[v^j] + (e^j{}_\lambda e_k[e_i^\lambda] + \Gamma_\mu^\lambda{}^\sigma e_i^\sigma e_k^\mu e^j{}_\lambda) v^i \right) (\vartheta^k)_a (e_j)^b \quad (\text{C.6})$$

$$= (e_k[v^j] + A_k{}^j{}_i v^i) (\vartheta^k)_a (e_j)^b. \quad (\text{C.7})$$

The second line follows from plugging in Equations (C.1)-(C.3) and applying product rule.

The third line follows from distributivity. In the fourth line, we simply *define* the $A_k{}^j{}_i$ coefficients:

$$A_k{}^j{}_i := e^j{}_\lambda e_k[e_i^\lambda] + \Gamma_\mu^\lambda{}^\sigma e_i^\sigma e_k^\mu e^j{}_\lambda. \quad (\text{C.8})$$

The derivative of a matrix can be related to the derivative of its inverse as follows:

$$e_k[e_i^\lambda] = -e_\ell^\lambda e_i^\mu e_k[e_\mu^\ell]. \quad (\text{C.9})$$

Plugging this into Equation (C.8), we can also write the $A_k^{j_i}$ coefficients in terms of the Christoffel symbols as follows:

$$A_k^{j_i} = -e_i^\mu e_k[e_\mu^j] + \Gamma_\mu^\lambda{}_\sigma e_i^\sigma e_k^\mu e_\lambda^j. \quad (\text{C.10})$$

The analogous computation can also be carried out for covectors and higher order tensors.

This gives us Equation (4.35):

$$\begin{aligned} \nabla_\mu T^{\sigma_1 \dots \sigma_r}{}_{\rho_1 \dots \rho_s} &= e_\mu [T^{\sigma_1 \dots \sigma_r}{}_{\rho_1 \dots \rho_s}] \\ &\quad + \sum_{i=1}^r A_\mu^{\sigma_i}{}_\lambda T^{\sigma_1 \dots \lambda \dots \sigma_r}{}_{\rho_1 \dots \rho_s} \\ &\quad - \sum_{j=1}^s A_\mu^\lambda{}_{\rho_j} T^{\sigma_1 \dots \sigma_r}{}_{\rho_1 \dots \lambda \dots \rho_s} \end{aligned}$$

Let us plug the basis vectors themselves into the expression above (Equation (4.35)). For each basis vector $(e_i)^a$, the coordinates functions are constants (a value of 1 for the i -th coordinate function and identically zero for all others). Therefore, we get the following:

$$\nabla_k (e_i)^b = A_k^\ell{}_m \delta^{im} (e_\ell)^b \quad (\text{C.11})$$

$$= A_k^\ell{}_i (e_\ell)^b \quad (\text{C.12})$$

By contracting both sides against the j -th dual basis, we can relate the $A_k^j{}_i$ coefficients to the spin connection coefficients, which are defined in terms of how the basis vectors change:

$$\omega_k^j{}_i := (\vartheta^j)_b \nabla_k (e_i)^b \quad (\text{C.13})$$

$$= A_k^\ell{}_i (e_\ell)^b (\vartheta^j)_b \quad (\text{C.14})$$

$$= A_k^\ell{}_i \delta_\ell^j \quad (\text{C.15})$$

$$= A_k^j{}_i \quad (\text{C.16})$$

We see, then, that the spin connection coefficients not only describe how the frame changes as we move around on the manifold, but also characterize how the covariant derivative deviates from a simple directional derivative. This makes intuitive sense since a “geometric” derivative, when computing how a tensorial quantity changes, must discount the change in the frame itself.

C.2 Structure Equations for Surfaces

For an orthonormal frame field adapted to a surface (i.e., a frame field where two of the basis vectors always lie tangent to a surface), the connection coefficient values $\omega_a^\mu{}_\nu$ are not allowed to be arbitrary. Clearly there are antisymmetry properties (under exchange of the last two indices) that come from the orthonormality constraint itself. But there are also integrability conditions that must be satisfied for a legitimate surface.

Define the following matrices of 1-forms (alternatively, matrix valued 1-forms):

$$\underline{\omega} := \begin{bmatrix} \omega_a^1{}_1 & \omega_a^1{}_2 & \omega_a^1{}_3 \\ \omega_a^2{}_1 & \omega_a^2{}_2 & \omega_a^2{}_3 \\ \omega_a^3{}_1 & \omega_a^3{}_2 & \omega_a^3{}_3 \end{bmatrix} \quad (\text{C.17})$$

$$\underline{\vartheta} := \begin{bmatrix} (\vartheta^1)_b \\ (\vartheta^2)_b \\ (\vartheta^3)_b \end{bmatrix} \quad (\text{C.18})$$

Then also define the wedge product between two matrices of differential forms to follow the standard matrix multiplication procedure, but replace the multiplication of individual elements by the wedge product of the elements. Likewise, define the exterior derivative of a matrix to be the matrix of element-wise exterior differentiations. With these notational conventions, we can succinctly describe the integrability conditions for a surface.

Theorem 2. *In an orthonormal frame:*

$$d\underline{\vartheta} = -\underline{\omega} \wedge \underline{\vartheta} \quad (\text{C.19})$$

$$d\underline{\omega} = -\underline{\omega} \wedge \underline{\omega} \quad (\text{C.20})$$

Proof of Theorem 2: For simplicity, let us arrange some more of 1-forms into matrices.

We first form a matrix of the standard dual basis 1-forms in \mathbb{R}^3 :

$$\underline{dx} := \begin{bmatrix} (dx^1)_b \\ (dx^2)_b \\ (dx^3)_b \end{bmatrix}. \quad (\text{C.21})$$

Let us also define a rotation matrix \underline{A} (each element is a scalar function, or 0-form):

$$\underline{A} := A_\mu{}^\nu = \begin{bmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{bmatrix}, \quad (\text{C.22})$$

such that

$$\underline{\vartheta} = \underline{A} \wedge \underline{dx}. \quad (\text{C.23})$$

This is possible because the standard basis is orthonormal, as is our orthonormal frame.

(For convenience, we'll also always choose our frame to be right-handed.)

We now show that $\underline{\omega} = \underline{A} \wedge d\underline{A}^T$. This will then make the proof to both parts of

the theorem quite straightforward.

$$\omega_a^\mu{}_\nu = (\vartheta^\mu)_b \nabla_a (e_\nu)^b \quad (\text{C.24})$$

$$= (\vartheta^\mu)_b g^{bc} \nabla_a (\vartheta^\nu)_c \quad (\text{C.25})$$

$$= \left(A_\lambda^\mu (dx^\lambda)_b \right) g^{bc} \nabla_a (A_\sigma^\nu (dx^\sigma)_c) \quad (\text{C.26})$$

$$= A_\lambda^\mu g^{bc} (dx^\lambda)_b (dx^\sigma)_c \nabla_a (A_\sigma^\nu) \quad (\text{C.27})$$

$$= A_\lambda^\mu \delta^{\lambda\sigma} \nabla_a (A_\sigma^\nu) \quad (\text{C.28})$$

Going from the first line to the second, we use metric compatibility of the covariant derivative. To get the third line, we write our orthonormal dual basis in terms of the standard dual basis. For the fourth line, we use the fact that the standard dual basis does not change spatially. Finally we use the orthonormality of the dual basis to get the final line. This proves the claim above. We now prove Equation (C.19):

$$\underline{\vartheta} = \underline{A} \wedge \underline{dx} \quad (\text{C.29})$$

$$d\underline{\vartheta} = d\underline{A} \wedge \underline{dx} + 0 \quad (\text{C.30})$$

$$= d\underline{A} \wedge (\underline{A}^T \wedge \underline{A}) \wedge \underline{dx} \quad (\text{C.31})$$

$$= (d\underline{A} \wedge \underline{A}^T) \wedge (\underline{A} \wedge \underline{dx}) \quad (\text{C.32})$$

$$= \underline{\omega}^T \wedge \underline{\vartheta} \quad (\text{C.33})$$

$$= -\underline{\omega} \wedge \underline{\vartheta} \quad (\text{C.34})$$

And now we prove Equation (C.20):

$$\underline{\omega} = \underline{A} \wedge d\underline{A}^T \quad (\text{C.35})$$

$$d\underline{\omega} = d\underline{A} \wedge d\underline{A}^T \quad (\text{C.36})$$

$$= d\underline{A} \wedge (\underline{A}^T \wedge \underline{A}) \wedge d\underline{A}^T \quad (\text{C.37})$$

$$= \underline{\omega}^T \wedge \underline{\omega} \quad (\text{C.38})$$

$$= -\underline{\omega} \wedge \underline{\omega} \quad (\text{C.39})$$

■

Corollary 1. *In an orthonormal frame:*

$$d_a(\vartheta^1)_b = \omega_a^2{}_1 \wedge (\vartheta^2)_b + \omega_a^3{}_1 \wedge (\vartheta^3)_b \quad (\text{C.40})$$

$$d_a(\vartheta^2)_b = \omega_a^1{}_2 \wedge (\vartheta^1)_b + \omega_a^3{}_2 \wedge (\vartheta^3)_b \quad (\text{C.41})$$

$$0 = \omega_a^1{}_3 \wedge (\vartheta^1)_b + \omega_a^2{}_3 \wedge (\vartheta^2)_b \quad (\text{C.42})$$

$$d_a \omega_b^2{}_1 = \omega_a^3{}_1 \wedge \omega_b^2{}_3 \quad (\text{C.43})$$

$$d_a \omega_b^3{}_1 = \omega_a^2{}_1 \wedge \omega_b^3{}_2 \quad (\text{C.44})$$

$$d_a \omega_b^3{}_2 = \omega_a^1{}_2 \wedge \omega_b^3{}_1 \quad (\text{C.45})$$

If these differential forms are evaluated only on vectors tangent to the surface (which is the only case we are interested in), then the second terms on the right-hand side of Equations (C.40) and (C.41) vanish.

C.3 Theorema Egregium

Equation (C.43) can be used to prove the famous Theorem of Gauss.

Theorem 3. (*Theorema Egregium*) *Gaussian curvature K is an intrinsic quantity.*

Proof of Theorem 2: Let us assume as usual we have an orthonormal frame where the first two basis vectors are tangent to the surface. Note that the left-hand side of Equation (C.43) is an anti-symmetric 2-form. This anti-symmetric 2-form is determined by its action on the tensors $(e_1)^a \otimes (e_2)^b$, $(e_2)^a \otimes (e_3)^b$, and $(e_3)^a \otimes (e_1)^b$. By plugging into the right-hand side of Equation (C.43), we see that the 2-form must be a scaling of $(\vartheta^1)_a \wedge (\vartheta^2)_b$. Furthermore, we see that the action on $(e_1)^a \otimes (e_2)^b$ simply computes the negative determinant of the Shape operator. So:

$$d_a \omega_b^2{}_1 = -K (\vartheta^1)_a \wedge (\vartheta^2)_b \quad (\text{C.46})$$

Notice that $\omega_a^2{}_1$ does not involve any extrinsic information. Additionally, the 3D exterior derivative acting on $\omega_a^2{}_1$ outputs a 2-form that also lives on the surface. The 1-form $\omega_a^2{}_1$ can then be replaced with its pull-back to the surface (via the surface embedding) and the 3D exterior derivative can be replaced with the corresponding surface exterior derivative. This means that the Gaussian curvature K is computable from intrinsic information only. ■

Appendix D

Image and Surface Curves

D.1 Basic Facts

In this section, we derive some basic results that will be used for proving later statements. These include computations of the connection coefficients for the radial frame. Along the way, we show many of the intermediate steps to illustrate how the mathematical formalism of Chapter 5 can be used in practice. (Refer to Chapter 5 for the definitions of many of the quantities.)

Fact 1. $\nabla_b V^a = -\delta_b^a$

Here, δ_b^a is a $(1, 1)$ tensor representing the identity map. In a fixed basis for the domain and range, its components are simply δ_ν^μ .

Proof of Fact 1:

$$\nabla_b V^a = \nabla_b (C - p)^a \quad (\text{D.1})$$

$$= 0 - \delta_b^a \quad (\text{D.2})$$

The second line follows since C is a constant position that does not depend on p , and the location of point p is by definition the quantity being varied. ■

Fact 2. $\nabla_a ||V||_g = -\hat{v}_a$

Proof of Fact 2:

$$\nabla_a ||V||_g = \nabla_a (V_b V^b)^{1/2} \quad (\text{D.3})$$

$$= \frac{1}{2} \frac{1}{||V||_g} 2V_b \nabla_a V^b \quad (\text{D.4})$$

$$= \frac{1}{||V||_g} V_b (-\delta_a^b) \quad (\text{D.5})$$

$$= -\frac{V_a}{||V||_g} \quad (\text{D.6})$$

$$= -\hat{v}_a \quad (\text{D.7})$$

■

Fact 3. $\nabla_a \hat{v}_b = \frac{1}{||V||_g} (-\delta_{ab} + \hat{v}_a \hat{v}_b)$

Proof of Fact 3:

$$\nabla_a \hat{v}_b = \nabla_a \left(\frac{V_b}{\|V\|_g} \right) \quad (\text{D.8})$$

$$= \frac{\nabla_a V_b}{\|V\|_g} - \frac{V_b}{\|V\|_g^2} \nabla_a \|V\|_g \quad (\text{D.9})$$

$$= \frac{1}{\|V\|_g} (-\delta_{ab} - \hat{v}_b \nabla_a \|V\|_g) \quad (\text{D.10})$$

$$= \frac{1}{\|V\|_g} (-\delta_{ab} + \hat{v}_a \hat{v}_b) \quad (\text{D.11})$$

■

Claim 6. $\nabla_a(\hat{n}^b \hat{v}_b) = -(\kappa_r \hat{r}_a + \tau_r \hat{r}_{\perp a}) \sin \alpha + \frac{1}{\|V\|_g} (-\hat{n}_a + \hat{v}_a \cos \alpha)$

Proof of Claim 6:

$$\nabla_a(\hat{n}^b \hat{v}_b) = -\Pi_a^b \hat{v}_b + \hat{n}^b (\nabla_a \hat{v}_b) \quad (\text{D.12})$$

$$= -\Pi_a^b P_b^c \hat{v}_c + \hat{n}^b \frac{1}{\|V\|_g} (-\delta_{ab} + \hat{v}_a \hat{v}_b) \quad (\text{D.13})$$

$$= -\Pi_a^b \hat{r}_b \sin \alpha + \frac{1}{\|V\|_g} (-\hat{n}_a + \hat{v}_a \cos \alpha) \quad (\text{D.14})$$

$$= -(\kappa_r \hat{r}_a + \tau_r \hat{r}_{\perp a}) \sin \alpha + \frac{1}{\|V\|_g} (-\hat{n}_a + \hat{v}_a \cos \alpha) \quad (\text{D.15})$$

■

Corollary 2. For orthogonal view projection, $\nabla_a(\hat{n}^b \hat{v}_b) = -(\kappa_r \hat{r}_a + \tau_r \hat{r}_{\perp a}) \sin \alpha$. For directional lighting, $\nabla_a(\hat{n}^b \hat{\ell}_b) = -(\kappa_l \hat{l}_a + \tau_l \hat{l}_{\perp a}) \sin \theta$.

Fact 4. $\nabla_b r^a = -\delta_b^a + (\hat{n}^c V_c) \Pi_b^a + \Pi_b^c V_c \hat{n}^a + \hat{n}^a \hat{n}_b$

Proof of Fact 4:

$$\nabla_b r^a = \nabla_b (V^a - \hat{n}^a (\hat{n}^c V_c)) \quad (\text{D.16})$$

$$= -\delta_b^a - (\nabla_b \hat{n}^a) (\hat{n}^c V_c) - \hat{n}^a (\nabla_b \hat{n}^c) V_c - \hat{n}^a \hat{n}^c (\nabla_b V_c) \quad (\text{D.17})$$

$$= -\delta_b^a + (\hat{n}^c V_c) \Pi_b^a + \Pi_b^c V_c \hat{n}^a + \hat{n}^a \hat{n}_b \quad (\text{D.18})$$

■

Fact 5.

$$\nabla_b \hat{r}^a = \frac{-\delta_b^a + (\hat{n}^c V_c) \Pi_b^a + \Pi_b^d V_d \hat{n}^a + \hat{n}^a \hat{n}_b + \hat{r}^a \hat{r}_b - \hat{r}^a (\kappa_r \hat{r}_b + \tau_r \hat{r}_{\perp b}) (\hat{n}^e V_e)}{\|r\|_g}$$

Proof of Fact 5:

$$\nabla_b \hat{r}^a = \nabla_b \left(\frac{r^a}{(r_c r^c)^{1/2}} \right) \quad (\text{D.19})$$

$$= \frac{\nabla_b r^a}{\|r\|_g} - \frac{1}{2} \frac{r^a}{(r^d r_d)^{3/2}} (2r_c \nabla_b r^c) \quad (\text{D.20})$$

$$= \frac{\nabla_b r^a}{\|r\|_g} - \frac{r^a}{r^d r_d} (-\hat{r}_b + (\kappa_r \hat{r}_b + \tau_r \hat{r}_{\perp b}) (\hat{n}^c V_c)) \quad (\text{D.21})$$

$$= \frac{-\delta_b^a + (\hat{n}^c V_c) \Pi_b^a + \Pi_b^d V_d \hat{n}^a + \hat{n}^a \hat{n}_b + \hat{r}^a \hat{r}_b - \hat{r}^a (\kappa_r \hat{r}_b + \tau_r \hat{r}_{\perp b}) (\hat{n}^e V_e)}{\|r\|_g} \quad (\text{D.22})$$

■

Fact 6. $\hat{r}^b \nabla_b \hat{r}^a = \tau_r \cot \alpha \hat{r}_\perp^a + \kappa_r \hat{n}^a$.

Proof of Fact 6:

$$\hat{r}^b \nabla_b \hat{r}^a = \frac{-\hat{r}^a + (\kappa_r \hat{r}^a + \tau_r \hat{r}_\perp^a)(\hat{n}^c V_c) + (\kappa_r \hat{r}^d + \tau_r \hat{r}_\perp^d) V_d \hat{n}^a + \hat{r}^a - \kappa_r \hat{r}^a (\hat{n}^e V_e)}{\|r\|_g} \quad (\text{D.23})$$

$$= \tau_r \frac{\cos \alpha}{\sin \alpha} \hat{r}_\perp^a + \kappa_r \frac{\sin \alpha}{\sin \alpha} \hat{n}^a \quad (\text{D.24})$$

$$= \tau_r \cot \alpha \hat{r}_\perp^a + \kappa_r \hat{n}^a \quad (\text{D.25})$$

■

Fact 7. $\hat{r}_\perp^b \nabla_b \hat{r}^a = \frac{-\hat{r}_\perp^a}{\|r\|_g} + \bar{\kappa}_r \cot \alpha \hat{r}_\perp^a + \tau_r \hat{n}^a$.

Proof of Fact 7:

$$\hat{r}_\perp^b \nabla_b \hat{r}^a = \frac{-\hat{r}_\perp^a + (\tau_r \hat{r}^a + \bar{\kappa}_r \hat{r}_\perp^a)(\hat{n}^c V_c) + (\tau_r \hat{r}^d + \bar{\kappa}_r \hat{r}_\perp^d) V_d \hat{n}^a - \hat{r}_\perp^a \tau_r (\hat{n}^e V_e)}{\|r\|_g} \quad (\text{D.26})$$

$$= \frac{-\hat{r}_\perp^a}{\|r\|_g} + \bar{\kappa}_r \frac{\cos \alpha}{\sin \alpha} \hat{r}_\perp^a + \tau_r \frac{\sin \alpha}{\sin \alpha} \hat{n}^a \quad (\text{D.27})$$

$$= \frac{-\hat{r}_\perp^a}{\|r\|_g} + \bar{\kappa}_r \cot \alpha \hat{r}_\perp^a + \tau_r \hat{n}^a \quad (\text{D.28})$$

■

Corollary 3. For orthographic projection, $\hat{r}_\perp^b \nabla_b \hat{r}^a = \bar{\kappa}_r \cot \alpha \hat{r}_\perp^a + \tau_r \hat{n}^a$.

By antisymmetry of the connection coefficients, the connection coefficients are then completely determined. Under an orthographic view projection, the connection coefficients are given by Equations (5.15)-(5.18).

Fact 8. *Let $\bar{v}^{\bar{a}}$ and $\bar{w}^{\bar{a}}$ be the eigenvectors of the illumination hessian, with corresponding eigenvalues λ and λ_{\perp} . Without loss of generality, let us take the ordered pair $(\bar{v}^{\bar{a}}, \bar{w}^{\bar{a}})$ to be right-handed. Then:*

$$\bar{v}^{\bar{a}} \check{H}_{\bar{a}}^{\bar{b}} = \lambda \bar{v}^{\bar{b}} \quad (\text{D.29})$$

$$\bar{w}^{\bar{a}} \check{H}_{\bar{a}}^{\bar{b}} = \lambda_{\perp} \bar{w}^{\bar{b}} \quad (\text{D.30})$$

$$\bar{v}^{\bar{a}} (\check{\nabla}_{\bar{a}} \check{\xi}_{\perp}^{\bar{b}}) = \lambda \bar{w}^{\bar{b}} \quad (\text{D.31})$$

$$\bar{w}^{\bar{a}} (\check{\nabla}_{\bar{a}} \check{\xi}_{\perp}^{\bar{b}}) = -\lambda_{\perp} \bar{v}^{\bar{b}} \quad (\text{D.32})$$

$$\check{g}_{\bar{a}\bar{c}} \bar{v}^{\bar{a}} (\check{\nabla}_{\bar{b}} \check{\xi}_{\perp}^{\bar{c}}) = -\lambda_{\perp} \check{g}_{\bar{d}\bar{b}} \bar{w}^{\bar{d}} \quad (\text{D.33})$$

$$\check{g}_{\bar{a}\bar{c}} \bar{w}^{\bar{a}} (\check{\nabla}_{\bar{b}} \check{\xi}_{\perp}^{\bar{c}}) = \lambda \check{g}_{\bar{d}\bar{b}} \bar{v}^{\bar{d}} \quad (\text{D.34})$$

Proof Fact 8: Equations (D.29) and (D.30) follow from the definition of $\bar{v}^{\bar{a}}$ and $\bar{w}^{\bar{a}}$ as eigenvectors. We will prove the remaining relations in the standard coordinate basis for the plane. Since the quantities we are computing are tensorial, they do not depend on the particular basis we use to compute them. In the standard coordinates, let us define the 90 degree right-handed (counter-clockwise) rotation matrix $\bar{A}_{\bar{\mu}}^{\bar{\nu}}$ by:

$$\bar{A}_{\bar{\mu}}^{\bar{\nu}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (\text{D.35})$$

Then we can relate the illumination hessian to the Jacobian of the isophote components as:

$$\check{\nabla}_{\check{\sigma}} \check{\xi}_{\perp}^{\check{\nu}} = \bar{A}_{\bar{\mu}}^{\bar{\nu}} \check{H}_{\check{\sigma}}^{\check{\mu}} \quad (\text{D.36})$$

This relation leads us to Equations (D.31) and (D.32) by contracting Equation (D.36) with $\bar{\nu}^{\check{\sigma}}$ and $\bar{w}^{\check{\sigma}}$ respectively. To prove Equations (D.33) and (D.34), we must contract our eigenvectors against the other axis of $\check{\nabla}_{\check{\sigma}} \check{\xi}_{\perp}^{\check{\nu}}$. Looking at the right-hand side of Equation (D.36), this means we contract first with the transposition of $\bar{A}_{\bar{\mu}}^{\bar{\nu}}$ (which applies a 90 degree clockwise rotation) and then with the illumination hessian (which is unaffected by the transposition since it is symmetric). This yields Equations (D.33) and (D.34). ■

D.2 Screen Coordinate Vector Fields

In this section, we seek to solve for vector fields that are tangent to a surface, and whose projections to the image plane remain constant. We can then use such vector fields to set up cartesian coordinates on the screen because their projections, as seen on the screen, do not rotate around. Let w^a be one such vector field. Since it is a vector tangent to the surface, we can write w^a as

$$w^a = A\hat{r}^a + B\hat{r}_{\perp}^a, \quad (\text{D.37})$$

where A and B are scalar functions and the component in the normal direction is identically zero. At a particular point of interest, the evaluations of A and B at that point will be known (we will want to make them match some desired basis on screen). Therefore, we just have to determine how the derivatives of A and B behave.

The constraint on derivatives comes from the desire that screen projections do not change:

$$0_c{}^a = \nabla_c(w^a - (w^b \hat{v}_b) \hat{v}^a) \quad (\text{D.38})$$

$$\Rightarrow 0_c{}^a = (\nabla_c w^a) - (\nabla_c w^b) \hat{v}_b \hat{v}^a. \quad (\text{D.39})$$

We made use of the fact that for orthographic projection, $\nabla_c \hat{v}^a$ is zero. Given Equation (D.37), we can write the derivative of w^a as

$$\begin{aligned} \nabla_c w^a &= (\nabla_c A) \hat{r}^a + (\nabla_c B) \hat{r}_\perp^a \\ &\quad + A(\omega_c{}^2{}_1 \hat{r}_\perp^a + \omega_c{}^3{}_1 \hat{n}^a) \\ &\quad + B(\omega_c{}^1{}_2 \hat{r}^a + \omega_c{}^3{}_2 \hat{n}^a) \end{aligned} \quad (\text{D.40})$$

$$\begin{aligned} &= (\nabla_c A + B \omega_c{}^1{}_2) \hat{r}^a \\ &\quad + (\nabla_c B + A \omega_c{}^2{}_1) \hat{r}_\perp^a \\ &\quad + (A \omega_c{}^3{}_1 + B \omega_c{}^3{}_2) \hat{n}^a. \end{aligned} \quad (\text{D.41})$$

Plugging Equation (D.41) into Equation (D.39), and contracting with \hat{r}_a and $\hat{r}_\perp{}_a$ yields the

following two relations (contracting against \hat{n}_a does not yield a distinct relation):

$$0 = \cos^2 \alpha (\nabla_c A + B \omega_c^1{}_2) - \cos \alpha \sin \alpha (A \omega_c^3{}_1 + B \omega_c^3{}_2) \quad (\text{D.42})$$

$$0 = \nabla_c B + A \omega_c^2{}_1 \quad (\text{D.43})$$

These in turn allow us to solve for the derivatives of the coefficients for w^a .

$$\nabla_c A = \tan \alpha (A \omega_c^3{}_1 + B \omega_c^3{}_2) - B \omega_c^1{}_2 \quad (\text{D.44})$$

$$\nabla_c B = -A \omega_c^2{}_1 \quad (\text{D.45})$$

D.3 Curve Relations

Here we prove the claims of Section 5.3.2, which make the following restrictions (for more details, refer to Section 5.3.2 itself):

- (1) Directional lighting and orthographic camera.
- (2) Miner's headlamp.
- (3) Critical points not coming from normal incidence.
- (4) Full curve zero set.

Claim 1 (*Intersection*) *Under restrictions (1)-(4), parabolic lines, suggestive contours, and Saint-Venant curves all intersect at critical points of illumination.*

Proof of Claim 1: We show that the defining equations for each curve type are satisfied at illumination minima.

- (1) Parabolic curves: Illumination critical points (besides those arising from normal incidence) are known to fall on parabolic curves (Koenderink [35]).
- (2) Suggestive contours: At a critical points of illumination, we have that for every surface vector $\vec{v}^{\bar{a}}$, $\vec{v}^{\bar{a}} \vec{\nabla}_{\bar{a}} \cos \theta = 0$. In particular, if we let $\vec{v}^{\bar{a}} := -\hat{r}^{\bar{a}}$, then we get:

$$\vec{v}^{\bar{a}} \vec{\nabla}_{\bar{a}} \cos \theta = \hat{r}^{\bar{a}} (\kappa_r \hat{r}_{\bar{a}} + \tau_r \hat{r}_{\perp \bar{a}}) \sin \theta \quad (\text{D.46})$$

$$= \kappa_r \sin \theta \quad (\text{D.47})$$

$$= 0 \quad (\text{D.48})$$

The quantity $\sin \theta$ is nonzero since the critical point does not arise from normal incidence. So $\kappa_r = 0$, and the point is also a suggestive contour point.

- (3) Saint-Venant curves: A critical point of illumination implies the image gradient $\vec{\xi}^{\bar{a}}$ is zero. So the Saint-Venant equation is trivially satisfied.

■

Claim 2 (Tangency) Under restrictions (1)-(4), the suggestive contour tangent lines up with parabolic curve tangent at critical points of illumination.

Proof of Claim 2: In our proof, we will make use of two facts:

Fact 9. (Koenderink [35]) Under miner's headlamp and orthogonal view projection, at a critical point of illumination on a parabolic curve, the light is in the plane spanned by normal \hat{n}^a and flat direction f^b .

Fact 10. Under miner's headlamp and orthogonal view projection, at a critical point of illumination on a parabolic curve, the radial direction is the flat direction.

We will prove the claim by writing down the differential equation for the curve tangent for each curve type. We will see that the defining equations are the same.

(1) For suggestive contours, we seek a vector direction t^c such that:

$$t^c \nabla_c (\Pi_{ab} \hat{r}^a \hat{r}^b) = 0 \quad (\text{D.49})$$

$$\Rightarrow t^c ((\nabla_c \Pi_{ab}) \hat{r}^a \hat{r}^b + 2\Pi_{ab} \hat{r}^a \nabla_c \hat{r}^b) = 0 \quad (\text{D.50})$$

$$\Rightarrow t^c (\nabla_c \Pi_{ab}) \hat{r}^a \hat{r}^b = 0 \quad (\text{D.51})$$

The third line follows from the second line by noting that the radial direction is the flat direction at the critical point of illumination, so $\Pi_{ab} \hat{r}^a = 0_b$.

(2) For parabolic curves, define \hat{f}^b to be the flat direction vector field on the parabolic curve. Off the parabolic curve, choose \hat{f}^b to be the principal direction that smoothly extends the vector field. We seek a vector direction t^c such that :

$$t^c \nabla_c (\Pi_{ab} \hat{f}^a \hat{f}^b) = 0 \quad (\text{D.52})$$

$$\Rightarrow t^c ((\nabla_c \Pi_{ab}) \hat{f}^a \hat{f}^b + 2\Pi_{ab} \hat{f}^a \nabla_c \hat{f}^b) = 0 \quad (\text{D.53})$$

$$\Rightarrow t^c (\nabla_c \Pi_{ab}) \hat{f}^a \hat{f}^b = 0 \quad (\text{D.54})$$

The third line follows from the second because at the critical point of illumination, the flat direction is simply \hat{r}^a . It does not matter how the vector field extends away from the illumination minimum. Since we arrive at the same differential equation, we also arrive at the same solution for t^c . ■

Claim 3 *Under restrictions (1)-(4), the tangent direction of Saint-Venant curves correspond to eigenvector directions of the illumination hessian, $\check{H}_a^{\bar{b}}$, at critical points of illumination.*

Proof of Claim 3: The Saint-Venant curve can be defined using the implicit function in Equation (5.40). At a critical point of illumination, the gradient of this implicit function is zero since we know the Saint-Venant zero set self-intersects at such points. This is verified as follows:

$$\check{\nabla}_{\bar{d}}(\bar{\xi}_{\perp}[\|\bar{\xi}\|_{\bar{g}}^2]) = \check{\nabla}_{\bar{d}}(2\check{g}_{\bar{b}\bar{c}}\bar{\xi}^{\bar{b}}\bar{\xi}_{\perp}^{\bar{a}}\check{H}_a^{\bar{c}}) \quad (\text{D.55})$$

$$= 2\check{g}_{\bar{b}\bar{c}}\left(\check{H}_d^{\bar{b}}\bar{\xi}_{\perp}^{\bar{a}}\check{H}_a^{\bar{c}} + \bar{\xi}^{\bar{b}}(\check{\nabla}_{\bar{d}}\bar{\xi}_{\perp}^{\bar{a}})\check{H}_a^{\bar{c}} + \bar{\xi}^{\bar{b}}\bar{\xi}_{\perp}^{\bar{a}}(\check{\nabla}_{\bar{d}}\check{H}_a^{\bar{c}})\right) \quad (\text{D.56})$$

$$= 0_{\bar{d}}. \quad (\text{D.57})$$

The last line follows from evaluating the second line at an illumination minimum where $\bar{\xi}^{\bar{a}} = 0^{\bar{a}}$ and $\bar{\xi}_{\perp}^{\bar{a}} = 0^{\bar{a}}$. Since the implicit function gradient is zero, we need to look at a higher derivative of the implicit function to determine the tangent directions. At an illumi-

nation minimum, the non-vanishing terms of implicit function hessian are:

$$\check{\nabla}_{\bar{e}}\check{\nabla}_{\bar{d}}(\bar{\xi}_{\perp}[\|\bar{\xi}\|_{\bar{g}}^2]) = 2\check{g}_{\bar{b}\bar{c}}\left(\check{H}_{\bar{d}}^{\bar{b}}(\check{\nabla}_{\bar{e}}\bar{\xi}_{\perp}^{\bar{a}})\check{H}_{\bar{a}}^{\bar{c}} + \check{H}_{\bar{e}}^{\bar{b}}(\check{\nabla}_{\bar{d}}\bar{\xi}_{\perp}^{\bar{a}})\check{H}_{\bar{a}}^{\bar{c}}\right). \quad (\text{D.58})$$

Tangents to the Saint-Venant curves are vectors $\bar{r}^{\bar{a}}$ such that

$$\bar{r}^{\bar{e}}\bar{r}^{\bar{d}}\check{\nabla}_{\bar{e}}\check{\nabla}_{\bar{d}}(\bar{\xi}_{\perp}[\|\bar{\xi}\|_{\bar{g}}^2]) = 0. \quad (\text{D.59})$$

If the implicit function hessian has negative determinant, we can solve for the two tangent directions. Equation (D.59) states that the tangent directions are simply the asymptotic directions of the implicit function hessian. If the implicit function hessian has positive determinant, then there is no solution. We will demonstrate two solutions below, showing that the implicit function hessian does not have positive determinant. In the process of determining the asymptotic directions, we will also argue that the implicit function hessian is generically non-singular, so we need not worry about having to take higher order derivatives to determine the tangent directions.

Let $\bar{v}^{\bar{a}}$ and $\bar{w}^{\bar{a}}$ be the eigenvectors of the *illumination* hessian, with eigenvalues λ and λ_{\perp} . Contracting each with the implicit function hessian given in Equation (D.58), we get (via Fact 8):

$$\bar{v}^{\bar{d}}\check{\nabla}_{\bar{e}}\check{\nabla}_{\bar{d}}(\bar{\xi}_{\perp}[\|\bar{\xi}\|_{\bar{g}}^2]) = 2\check{g}_{\bar{a}\bar{e}}(\lambda\lambda_{\perp}^2 - \lambda^2\lambda_{\perp})\bar{w}^{\bar{a}} \quad (\text{D.60})$$

$$\bar{w}^{\bar{d}}\check{\nabla}_{\bar{e}}\check{\nabla}_{\bar{d}}(\bar{\xi}_{\perp}[\|\bar{\xi}\|_{\bar{g}}^2]) = 2\check{g}_{\bar{a}\bar{e}}(\lambda\lambda_{\perp}^2 - \lambda^2\lambda_{\perp})\bar{v}^{\bar{a}} \quad (\text{D.61})$$

Since $\bar{v}^{\bar{a}}$ and $\bar{w}^{\bar{a}}$ are orthogonal with respect to the screen metric, these vectors are asymp-

otic directions for the implicit function hessian. The fact that we have two non-zero solutions means that the implicit function hessian does not have positive determinant. We furthermore note that as long as $\lambda \neq 0$, $\lambda_{\perp} \neq 0$, and $\lambda \neq \lambda_{\perp}$, the implicit function hessian is also non-singular. So the eigenvectors of the illumination hessian are indeed the tangent directions of the Saint-Venant curves at illumination minima. ■

Claim 4 *Under restrictions (1)-(4), the Saint-Venant curves are tangent to the suggestive and parabolic curves at critical points of illumination if and only if they all run perpendicular to the flat direction.*

Proof of Claim 4: Let us first relate the illumination hessian to information about the surface. We do so by taking advantage of the fact that at a critical point of illumination, the illumination hessian can be defined using any covariant derivative. This is true because, when applying the covariant derivative for the second time, all terms involving the connection coefficients are killed by the zero gradient. In particular, we can define the illumination hessian using the covariant derivative compatible with the surface metric.

$$\check{\nabla}_{\bar{a}}\check{\nabla}_{\bar{b}}\cos\theta = \bar{\nabla}_{\bar{a}}\bar{\nabla}_{\bar{b}}\cos\theta \quad (\text{D.62})$$

$$= -\bar{\nabla}_{\bar{a}}(\bar{\Pi}_{\bar{b}\bar{c}}\hat{r}^{\bar{c}}\sin\theta) \quad (\text{D.63})$$

$$= -(\bar{\nabla}_{\bar{a}}\bar{\Pi}_{\bar{b}\bar{c}})\hat{r}^{\bar{c}}\sin\theta - \bar{\Pi}_{\bar{b}\bar{c}}\bar{\nabla}_{\bar{a}}(\hat{r}^{\bar{c}}\sin\theta) \quad (\text{D.64})$$

The tangent of the Saint-Venant curve is an eigenvector of the illumination hessian. Let $\bar{v}^{\bar{a}}$ be an eigenvector with eigenvalue λ . Then plugging $\bar{v}^{\bar{a}}$ into the differential

equation for the suggestive contour and parabolic curve tangents given by Equation (D.49), we get:

$$0 = \bar{v}^{\bar{a}}(\bar{\nabla}_{\bar{a}}\bar{\Pi}_{\bar{b}\bar{c}})\hat{r}^{\bar{b}}\hat{r}^{\bar{c}} \quad (\text{D.65})$$

$$\Rightarrow 0 = -\frac{\bar{v}^{\bar{a}}}{\sin\theta} \left((\check{\nabla}_{\bar{a}}\check{\nabla}_{\bar{b}}\cos\theta)\hat{r}^{\bar{b}} + (\hat{r}^{\bar{b}}\bar{\Pi}_{\bar{b}\bar{c}})\bar{\nabla}_{\bar{a}}(\hat{r}^{\bar{c}}\sin\theta) \right) \quad (\text{D.66})$$

$$\Rightarrow 0 = -\frac{\bar{v}^{\bar{a}}}{\sin\theta}(\check{\nabla}_{\bar{a}}\check{\nabla}_{\bar{b}}\cos\theta)\hat{r}^{\bar{b}} \quad (\text{D.67})$$

$$\Rightarrow 0 = -\frac{\lambda}{\sin\theta}\check{g}_{\bar{b}\bar{c}}\check{v}^{\bar{c}}\hat{r}^{\bar{b}} \quad (\text{D.68})$$

The third line follows from the second line because the radial direction is also the flat direction at an illumination minimum.

So the Saint-Venant curve is tangent to the suggestive contour and parabolic curve at illumination minima if and only if all the curves are perpendicular to the radial direction. At illumination minima, the radial direction is also the flat direction, and being perpendicular to the radial direction on screen also corresponds to being perpendicular to the radial direction on the surface. ■

Claim 5 *Under restrictions (1)-(4), a suggestive contour point falls on a minimal curve if and only if the radial direction is 45 degrees (on the surface) from a principal direction.*

Proof of Claim 5: Let us work in the principal frame. In this frame, the Shape operator is diagonal and has entries λ and $-\lambda$. Let ϕ denote the angle (measured using the surface

metric) of the radial direction with the first principal direction. Then:

$$\kappa_r = \Pi_{ab} \hat{r}^a \hat{r}^b \quad (\text{D.69})$$

$$= \lambda \cos^2 \phi - \lambda \sin^2 \phi \quad (\text{D.70})$$

$$= \lambda (\cos \phi + \sin \phi)(\cos \phi - \sin \phi) \quad (\text{D.71})$$

So κ_r is zero if and only if, λ is zero or ϕ is 45 degrees. ■

D.4 Principal Directions

So far, in three dimensional space, we have related everything using the radial frame. This is convenient, although it may be desirable to relate the radial frame to the directions of maximal and minimal curvature. Here we present some such relations between the radial frame and the principal frame. Let κ_1 and κ_2 be the two eigenvalues of the shape operator. Then there is a choice of labeling these eigenvalues such that the (un-normalized) principal directions given in the radial basis are:

$$\tau_r \hat{r}^a + (\kappa_1 - \kappa_r) \hat{r}_\perp^a \quad \text{and} \quad (\kappa_2 - \bar{\kappa}_r) \hat{r}^a + \tau_r \hat{r}_\perp^a \quad (\text{D.72})$$

On a suggestive contour, these become:

$$\sqrt{-K} \hat{r}^a + \kappa_1 \hat{r}_\perp^a \quad \text{and} \quad -\kappa_1 \hat{r}^a + \sqrt{-K} \hat{r}_\perp^a \quad (\text{D.73})$$

We can also write down a similar expression relating the radial frame to the apparent curvature directions used for defining the apparent ridges on a surface. Let κ'_1 and κ'_2 be the two eigenvalues of $\check{g}^{ab}\mathbf{\Pi}_{bc}$. Then the (un-normalized) apparent principal directions given in the radial basis are (for some choice of labeling the eigenvalues):

$$\frac{\tau_r}{\cos^2 \alpha} \hat{r}^a + \left(\kappa'_1 - \frac{\kappa_r}{\cos^2 \alpha} \right) \hat{r}_\perp^a \quad \text{and} \quad (\kappa'_2 - \bar{\kappa}_r) \hat{r}^a + \tau_r \hat{r}_\perp^a \quad (\text{D.74})$$

Unlike the principal directions above, these eigenvectors are not generally orthogonal on the surface.

D.5 Apparent Curvature

Besides intersections and tangencies, another detectable piece of information we have in the image is the apparent curvature of a curve. Apparent curvature is (up to sign) the curvature of a curve as measured on the image screen. However, unlike the usual definition of a curve's curvature, which is always non-negative, we will take apparent curvature to be signed. We adopt the convention that the projected (outward pointing) normal of a surface and the projected curve tangent (in that order) form a right-handed basis for the screen. When the frame turns in a right-handed direction, curvature is deemed to be positive; if it rotates in a left-handed direction, it is deemed to be negative. Figure D.1 shows the choice of convention for a silhouette curve.

Notationally, we will use κ without a subscript for curve curvatures; κ with a subscript (e.g., κ_v) will denote normal curvatures. For some three-dimensional curve with

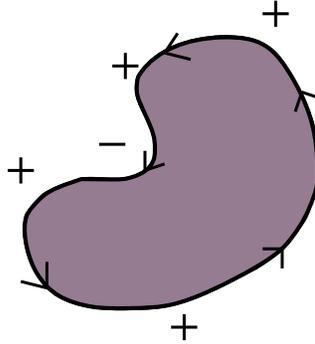


Figure D.1: Sign of apparent curvature of silhouette curve. As the curve is traced so that the object is on the left, twisting of the (normal, tangent) frame in the right-handed direction is positive; twisting in the other direction is negative.

curvature κ , it is not hard to show (by writing out equations in the standard coordinates for \mathbb{R}^3 and aligning the projection direction with one of the standard axes) that under orthographic projection the corresponding curve produced by screen-orthographic projection has apparent curvature $\check{\kappa}$ given by

$$\check{\kappa} = \kappa \frac{\cos \beta}{\cos^3 \varphi}, \quad (\text{D.75})$$

where β is the angle between the curve binormal and the projection direction, and φ is the angle (in 3D) between the curve tangent and its screen-orthographic projection on screen.

For two conjugate directions u^a and v^a (with angle ψ between them), the Gaussian curvature is related to the normal curvatures associated with each direction, κ_u and κ_v , by

$$K \sin^2 \psi = \kappa_u \kappa_v. \quad (\text{D.76})$$

Applied to silhouettes (where the tangent to the silhouette is conjugate to the radial direction), we get Koenderink's theorem [34]: on a silhouette curve, the apparent curvature of the silhouette curve is related to the Gaussian curvature and radial curvature of the surface by

$$K = -\check{\kappa}\kappa_r. \quad (\text{D.77})$$

Since κ_r is always non-positive (otherwise such a point would not lie on the silhouette), we see that the sign of the Gaussian curvature matches that of the apparent curvature.

We would like to obtain some similar type of relationship between apparent curvature and geometric information for non-silhouette curves as well. Let us define a set of curves on the surface as the zero set of some energy functional, f . We can write the gradient of this functional as

$$\nabla_a f = A\hat{r}_a + B\hat{r}_{\perp a}, \quad (\text{D.78})$$

where A and B are some scalar functions and there is no normal component since we are not interested in movement off the surface. Since we are using an orthonormal frame, we can write a unit vector tangent to the surface curve as

$$\hat{t}^a := \frac{1}{\sqrt{A^2 + B^2}} (-B\hat{r}^a + A\hat{r}_{\perp}^a). \quad (\text{D.79})$$

Here we had a little freedom in choosing whether we rotate in the right-handed direction or left-handed direction. Without loss of generality, we will assume A is negative. If our above

rotation does not yield a negative, just switch to $\hat{t}^a \mapsto -\hat{t}^a$. This corresponds to choosing the settled upon sign convention for the orientation of the projected curve used to define apparent curvature.

With this unit tangent vector, we can carry out the computations for computing apparent curvature. The expression is not particularly simple. However, if we restrict attention to curves where the viewer sees the surface more edge-on, the expression can be simplified. In the miner's headlamp case, this corresponds to looking at regions where the shading of the surface is dark (e.g., valleys of illumination). Taking the limit in which $\cos \alpha \rightarrow 0$ and $\sin \alpha \rightarrow 1$ yields

$$\check{\kappa} \approx -\frac{\kappa_{\hat{t}}}{\sin^2 \varphi'}, \quad (\text{D.80})$$

where φ' is the angle between the tangent to the curve \hat{t}^a and the radial direction (which becomes the projection direction in the limit). Note that this agrees with our previously derived result for silhouettes. If we only cull terms second order and higher in $\cos \alpha$, then we have

$$\check{\kappa} \approx -\frac{\kappa_{\hat{t}} \sin \alpha}{\sin^2 \varphi'} - \frac{\cos \alpha}{|A|^3} (-B^2 \hat{r}[A] + AB \hat{r}_{\perp}[A] + AB \hat{r}[B] - A^2 \hat{r}_{\perp}[B]). \quad (\text{D.81})$$

Note that the sign of the apparent contour is opposite that of the normal curvature associated with the tangent of the curve. Also note that when the curve starts to run more parallel to the view direction, the magnitude of the apparent curvature will tend to grow large.

D.6 General Lighting

For simplicity, we have thus far restricted attention to the miner's headlamp case, in which the light's position is co-located with the view camera's pinhole (at a point at infinity). This is perhaps a reasonable simplification since the miner's headlamp tends not to cast large shadows and lambertian objects are arguably well understood as a set of curves in the image.

In this section, however, we briefly note how these techniques carry over to general directional lighting conditions. See Figure 5.2 for the setup. Now, in addition to the radial direction \hat{r}^a , we have a new tangent vector, \hat{l}^a , that is derived from projecting the light direction orthogonally onto the tangent plane and normalizing with respect to the surface metric. We also have its ninety degree right-handed rotation \hat{l}_\perp^a that, along with the surface normal, can be used to define a new right-handed orthonormal frame.

Introducing new degrees of freedom leads many of the expressions to become significantly more complicated. To simplify the notation, we define some new variables meant to draw analogies with the simpler expressions in the miner's headlamp case. These variable names are chosen only to draw upon the analogy, they do not really represent normal curvatures or geodesic torsions anymore.

$$\kappa_{rl} := \kappa_r \cos \phi + \tau_r \sin \phi = \Pi_{ab} \hat{r}^a \hat{l}^b \quad (\text{D.82})$$

$$\tau_{rl} := \tau_r \cos \phi + \bar{\kappa}_r \sin \phi = \Pi_{ab} \hat{r}_\perp^a \hat{l}^b \quad (\text{D.83})$$

$$\tau_{lr} := \tau_r \cos \phi - \kappa_r \sin \phi = \Pi_{ab} \hat{r}^a \hat{l}_\perp^b \quad (\text{D.84})$$

$$\bar{\kappa}_{rl} := \bar{\kappa}_r \cos \phi - \tau_r \sin \phi = \Pi_{ab} \hat{r}_\perp^a \hat{l}_\perp^b \quad (\text{D.85})$$

Some of the new recurring expressions involve derivatives of the various angles. The derivatives are summarized below:

$$\nabla_a \alpha = \kappa_r \hat{r}_a + \tau_r \hat{r}_{\perp a}, \quad (\text{D.86})$$

$$\nabla_a \theta = \kappa_{rl} \hat{r}_a + \tau_{rl} \hat{r}_{\perp a}, \quad (\text{D.87})$$

$$\nabla_a \phi = (\tau_{lr} \cot \theta - \tau_r \cot \alpha) \hat{r}_a + (\bar{\kappa}_{rl} \cot \theta - \bar{\kappa}_r \cot \alpha) \hat{r}_{\perp a} \quad (\text{D.88})$$

$$= \Pi_{ab} \hat{l}_{\perp}^b \cot \theta - \Pi_{ab} \hat{r}_{\perp}^b \cot \alpha. \quad (\text{D.89})$$

The derivative of the light frame with respect to radial frame directions is given by:

$$\hat{r}^a \nabla_a \hat{l}^b = \tau_{lr} \cot \theta \hat{l}_{\perp}^b + \kappa_{rl} \hat{n}^b, \quad (\text{D.90})$$

$$\hat{r}_{\perp}^a \nabla_a \hat{l}^b = \bar{\kappa}_{rl} \cot \theta \hat{l}_{\perp}^b + \tau_{rl} \hat{n}^b, \quad (\text{D.91})$$

$$\hat{r}^a \nabla_a \hat{l}_{\perp}^b = -\tau_{lr} \cot \theta \hat{l}^b + \tau_{lr} \hat{n}^b, \quad (\text{D.92})$$

$$\hat{r}_{\perp}^a \nabla_a \hat{l}_{\perp}^b = -\bar{\kappa}_{rl} \cot \theta \hat{l}^b + \bar{\kappa}_{rl} \hat{n}^b. \quad (\text{D.93})$$

With these relations, one can carry out many of the computations done before for the miner's headlamp case.